MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, Sept. 2010, p. 341–362 Vol. 74, No. 3
1092-2172/10/$12.00 doi:10.1128/MMBR.00009-10


# ABC Transporters Involved in Export of Cell Surface Glycoconjugates

Leslie Cuthbertson,†‡ Veronica Kos,‡ and Chris Whitfield*

*Department of Molecular & Cellular Biology, University of Guelph, Guelph, Ontario, Canada N1G 2W1*



## INTRODUCTION

In prokaryotes, glycosylated macromolecules (glycoconjugates) often form the point of first contact between the organism and its immediate environment. Surface glycoconjugates play major roles in the maintenance of cell envelope structure, protection against host immune defenses, signaling events, and the formation of biofilms. The structures of bacterial glycans show tremendous diversity in terms of component sugars, noncarbohydrate modifications, and linkage configurations. They are also attached to a broad range of molecules, including lipids, peptidoglycan, and proteins.

Despite the remarkable diversity in oligo- and polysaccharide structures and the manners in which they are presented on the cell surface, the number of fundamentally different pathways used in polysaccharide synthesis and export across the inner membrane is relatively limited. For polysaccharides or oligosaccharides synthesized directly from activated precursors consisting of nucleotide mono- or diphosphoglycoses, there are currently just three known export strategies that have been described in any detail (Fig. 1); these have been termed "Wzx/Wzy dependent," "ATP-binding cassette (ABC) transporter dependent," and "synthase dependent," based on characteristic components. These three general assembly-export strategies are not confined to the assembly of any one class of glycoconjugate. It is the steps that occur beyond synthesis and export (i.e., attachment of the glycan to a final acceptor molecule) that define the type of glycoconjugate that is formed. Most of our current understanding of these systems has been obtained in the context of lipopolysaccharide (LPS) O-antigen polysaccharide (O-PS) and capsular or exopolysaccharide (CPS or EPS) assembly, and these have been reviewed elsewhere (121, 173).

In the Wzx/Wzy-dependent and ABC transporter-dependent pathways, the polysaccharide is built on a lipid acceptor. The activated precursors are cytosolic sugar nucleotides, and the glycosyltransferase reactions that transfer sugars to the lipid carrier occur at the cytoplasmic face of the membrane. Ultimately, both of these pathways lead to the formation of a lipid-linked polysaccharide located outside the cytoplasmic membrane. However, the intervening steps between chain initiation and completion are quite different. In the Wzx/Wzy-dependent pathway, individual undecaprenol diphosphate (Und-PP)-linked polysaccharide repeat units are assembled and exported across the membrane by a transport process requiring a Wzx protein homolog. Wzx is considered a "flippase," but its precise mechanism of action is unknown. The newly exported lipid-linked repeat units then form the substrates for a polymerization reaction that requires a Wzy (putative polymerase) homolog, which extends the growing chain one repeat unit at a time at the periplasmic face of the cytoplasmic membrane. The polymerization process is controlled by a member of the PCP (polysaccharide copolymerase) family (32, 96). In contrast, polysaccharides assembled by ABC transporters are fully polymerized by sequential glycosyl transfer at the cytoplasmic face of the inner membrane. The glycan can be assembled as a Und-PP-linked intermediate, as is the case for most O-PSs (121). Alternatively, for some CPSs, it seems more likely that the acceptor is diacylglycerol phosphate (173). The completed molecule is then exported by

* Corresponding author. Mailing address: Department of Molecular & Cellular Biology, University of Guelph, Guelph, Ontario, Canada N1G 2W1. Phone: (519) 824-4120, ext. 53361. Fax: (519) 837-3273. E-mail: cwhitfie@uoguelph.ca.

† Present address: Department of Biochemistry and Biomedical Sciences, McMaster University, 1200 Main St. W, Hamilton, Ontario, Canada L8S 3Z5.

‡ L.C. and V.K. contributed equally to this work.

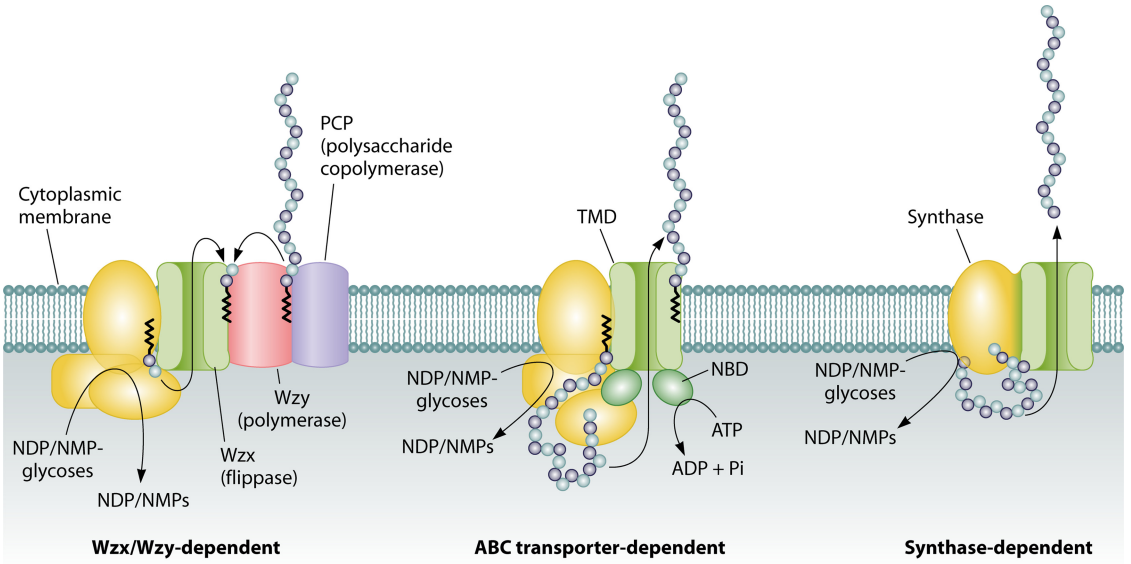


FIG. 1. Models for processes involved in the synthesis and *trans*-cytoplasmic membrane export of bacterial glycans. In the Wzx/Wzy-dependent pathway, individual lipid-linked repeat units are synthesized by glycosyltransferase enzymes (illustrated in yellow in each model) located at the interface of the cytoplasm and the membrane. The lipid-linked repeat units are exported via the Wzx flippase (a putative transporter of unknown mechanism) and polymerized at the periplasmic face of the membrane. The polymer grows in a blockwise process, by addition of new repeat units to the reducing end of the glycan in a reaction requiring the Wzy polymerase. In the ABC transporter-dependent pathway, the chain is elongated by addition of monomers to the nonreducing terminus of a lipid-linked intermediate and is completed in the cytoplasm, prior to export via the ABC transporter. In the synthase-dependent system, a single protein is thought to serve as both a polymerase and an exporter, but the details of the export process are unknown. There is no unifying involvement of a lipid acceptor in the synthase pathway, and even the direction of chain growth may differ, depending on the system. Glycans (or glycoses) derived from these pathways can be attached to protein, lipid, LPS, or peptidoglycan acceptors. In the case of some glycoconjugates in Gram-negative bacteria, the final cellular location may depend on additional export pathways to transfer the molecule across the outer membrane.

the ABC transporter (see below), but it is not clear whether, under normal physiological conditions, chain extension must be completed before export can begin. Components of the Wzx/Wzy-dependent and ABC transporter-dependent pathways have been identified in both Gram-positive and Gram-negative organisms, and some aspects of the underlying biochemistry have been described. Significantly less is known about a third assembly process, the synthase-dependent pathway (Fig. 1). Synthases are processive glycosyltransferases and are involved in the formation of some important biological molecules, including bacterial cellulose (130), hyaluronan and chondroitin (in some producing organisms [171]), alginate (123), and poly-β-D-*N*-acetylglucosamine (GlcNAc) (60). In the synthase pathway, the nature of the acceptor on which the polymer grows is not always certain, nor is the process by which the polymer (or repeat units) is exported. It is possible that a single protein (the synthase) is sufficient for both polymerization and export activities in some cases (66, 171), but details of the process are unknown.

Other export-assembly mechanisms may emerge as more information becomes available. One candidate includes those processes where the direct donors for glycosylation are polyprenol-linked glycoses and one or more residues are transferred to an acceptor at the periplasmic face of the cytoplasmic membrane. Although limited in diversity, such systems have been identified in phage-mediated glucosylation of LPS O-PSs (50, 74) and in the modification of LPS lipid A by 4-amino-4-deoxy-L-arabinose (4-aminoarabinose) (reviewed in reference 116). In these pathways, a single glycose residue is exported, and in each case, a Und-P-linked glycose is the donor. This process is superficially similar to the activity of Wzx, but the mechanism(s) of export for the Und-P-linked glycoses is unknown. Although candidate flippases have been identified, there is no obvious similarity to other glycotransporters, including Wzx. Glycosylation by a Und-P-linked donor is also a hallmark of the biosynthesis of the long-chain arabinan part of the mycolyl-arabinogalactan-peptidoglycan complex in mycobacteria (reviewed in references 13 and 154). The donor is decaprenol-monophospho-arabinose (Dec-P-Ara), whose synthesis occurs without the involvement of sugar nucleotides, but the cellular compartment in which its activity occurs, and hence any requirement for a specific exporter, is currently unknown (see below).

Members of the ABC transporter superfamily participate in a wide range of glycosylation processes. The involvement of ABC transporters in glycan export was first identified in studies of CPS biosynthesis in *Haemophilus influenzae* (79) and *Escherichia coli* (109, 143). In the absence of the ABC transporter, polymer is synthesized by each of these systems, but it never leaves the cytoplasm. Later, a similar phenomenon was recognized in the biosynthesis of LPS O-PS in *Yersinia enterocolitica* O:3 (182). Subsequently, ABC transporters were identified in the assembly of teichoic acid in *Bacillus subtilis* 168 (83) and in the protein glycosylation system from *Campylobacter jejuni* (2). ABC transporters are now known to be involved in the export of representatives from all of the major classes of cell surface glycoconjugates. By mining genomic information to examine the distribution of "glyco-focused" ABC transporters, it is clear

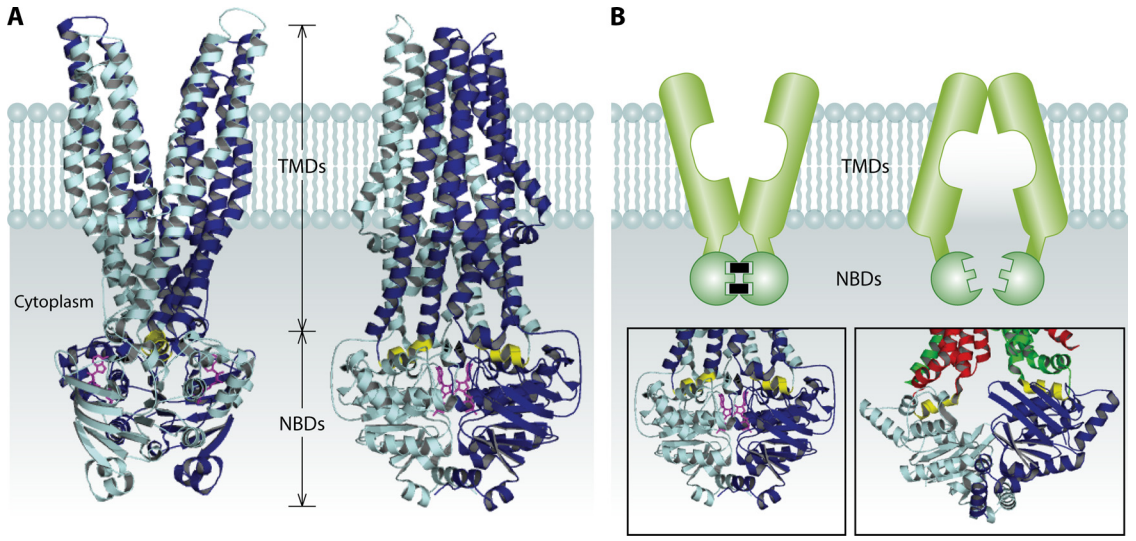


FIG. 2. Model for ABC exporters. (A) Structure of the complete Sav1866 multidrug exporter from *S. aureus*, illustrated from two angles rotated 90° around the vertical axis (35). Sav1866 is a half-transporter in which each subunit (identified by different shades of blue) contains one TMD fused to one NBD. The coupling helices that play a crucial role in opening the channel are highlighted in yellow. Although the Sav1866 crystal structure was obtained in the presence of ADP (magenta), its NBDs exhibit the sandwich dimer typically reserved for ATP-bound complexes (37, 144, 179). It has been suggested that during the crystallization process, the conformation of the transporter was shifted to the ATP-bound state (36). (B) Models of the open and closed conformations of ABC transporters. The model is based upon the information obtained from crystal structures of Sav1866 (open; outward facing) and ModABC (closed; inward facing), as described elsewhere (57). The model illustrates how ABC transporters may link ATP hydrolysis to distinct conformational changes in the transporter proteins. Conformational changes are transmitted from the NBDs with the hydrolysis of ATP to the TMDs through the interactions of the coupling helices (highlighted in yellow). ModBC is the molybdate/tungstate importer from *Archaeoglobus fulgidus* (57). The structural models were created from the structures under PDB accession numbers 2ONJ (Sav1866 with AMP-PNP) and 2ONK (ModBC), using PyMol (www.pymol.org/).

that they play pivotal roles in the glycobiology of many bacteria, even though there are many cases where the exact nature and structure of the export substrate have not been established. In each of these systems, the organism must overcome the same challenges. First, the export system must engage a biosynthesis pathway, often involving multiple proteins with different catalytic activities. Second, the exporter must handle a large hydrophilic (and sometimes charged) substrate, often attached to a hydrophobic lipid acceptor. Here we describe our current understanding of the structure and function of glycan ABC transporters and illustrate how bioinformatic analyses can extend this knowledge to other cell surface glycoconjugates.

## STRUCTURE AND FUNCTION OF ABC TRANSPORTERS

ABC transporters represent a massive protein superfamily and play diverse roles in biological processes which span the biological kingdoms. In bacteria, ABC transporters are involved in the uptake of nutrients (e.g., some sugars, amino acids, and vitamins). They also participate in the export of molecules with remarkable structural diversity, including various drugs and other small-molecule inhibitors, proteins, lipids, and (in the context of this review) oligo- and polysaccharides.

ABC transporters contain four essential domains: two transmembrane domains (TMDs), which are integral membrane proteins, and two nucleotide-binding domains (NBDs) that associate with the TMDs on the cytoplasmic face of the membrane (Fig. 2A and B). These domains can be organized as individual polypeptides or may be fused into multidomain proteins in a variety of formats. The TMDs of different transporters typically share low sequence similarity, and the number of α-helices varies, with 8 to 20 in importers and, typically, 12 in exporters (88, 97). The membrane-spanning α-helices of the TMDs form the transport channel, and the differences in the primary sequences of the TMDs may be a reflection of the wide variety of substrates used. In contrast to the sequence variability shared between the TMDs, the NBDs of ABC transporters possess the highly conserved sequences that define ABC proteins. The signature "LSGGQ motif" is used to identify members of the ABC transporter superfamily, and it resides between the Walker A and B motifs, which are found in a variety of ATP- and GTP-hydrolyzing proteins (167). ABC proteins also contain additional motifs (e.g., the Q and H loops) essential for ATP binding and hydrolysis (23, 136, 144, 179). Each NBD monomer forms two structural domains: a RecA-like domain, sharing structural similarities with other ATPases, and a smaller helical domain, which contains the signature motif. The ATP-binding sites are created at the interface of two NBDs, which align in the transporter in a head-to-tail arrangement (Fig. 2A and B). In some cases, the NBDs also contain accessory domains, which may aid in the transport of the substrate or be involved in regulatory mechanisms (14). A combination of structural information and biochemical evidence has led to proposals concerning the functions of the conserved ABC motifs in ATP hydrolysis and the mechanism by which ATP hydrolysis is coupled to substrate transport. In the working models, ATP binding induces conformational changes in the NBDs that force them into closer contact, creating the characteristic nucleotide sandwich (68). The ATP-

binding sites are created by the juxtaposition of the Walker A motif of one NBD and the signature motif on its partner. Structures of various NBDs in nucleotide-bound and unbound states clearly reveal significant changes in the proximity of the NBDs (Fig. 2B).

The bacterial multidrug exporter Sav1866 (35), the lipid A flippase MsbA (168), and the maltose transporter $MalFGK_2E$ (105) have become influential models for ABC transporters in both prokaryotes and eukaryotes. The Sav1866 exporter from *Staphylococcus aureus* is a typical "half-transporter," comprised of a homodimer in which each polypeptide contributes one TMD and one NBD (35) (Fig. 2A). The TMDs of the transporter consist of long helices that extend outward through the lipid bilayer. The TMDs interact with one another, forming two wings, with each one being composed of α-helices from each TMD subunit. The large pocket formed between the two TMDs in the crystal structure of Sav1866 is open to the periplasm and is thought to form the export cavity for the extruded antibiotic, as this part of the structure is not very hydrophobic and may therefore facilitate extrusion of hydrophobic substrates to the outer face of the membrane. The NBDs sit ~25 Å away from the inner membrane due to long loops in the TMDs, and each TMD is in contact with both NBDs (35). TMD-NBD contacts are mediated by coupling helices in each of the TMDs and through the Q loop and the TEVGERV sequence of the X loop in the NBDs. The X loop precedes the ABC signature motif in the NBDs and is unique to exporters. It is important for the interactions of the NBDs with the TMDs (35, 103). In each TMD subunit, one coupling helix makes contacts with both NBD subunits, while the other makes contacts only with the NBD from the opposite subunit. The arrangement of ABC transporters involved in the export of substrates is thought to resemble that of the Sav1866 transporter (35). However, the mechanisms by which transporters work may vary, as recent evidence suggests that BtuCD (the vitamin $B_{12}$ importer) has a gate conformation distinct from those of other importers, such as those specific for molybdate (ModBC) (46).

In the maltose ABC transporter, two polypeptides (MalFG) provide the TMDs, while the NBD dimer consists of two copies of MalK. The activity of the transporter is aided by a periplasmic binding protein, MalE, and the structure of the entire complex has been solved (105, 106). TMD-NBD interactions in the $MalFGK_2E$ complex are mediated by the EAA motifs in the MalF and MalG polypeptides (105). The EAA motif consists of two short helices. One of these is termed the coupling helix, but it is distinct from the coupling helices in Sav1866, and it interacts with the Q loop of a MalK monomer.

MsbA is another "half-transporter" and has been studied in detail (57, 136), with structures solved for multiple species and different conformational states (168). The varied conformational states are suggestive of a substantial structural change occurring during the catalytic cycle of the exporter, but this proposal remains to be confirmed *in vivo*. In the current working model, an ABC transporter cycles between the nucleotide-bound and unbound states. The separation of the NBDs is communicated to the TMDs via a coupling helix (Fig. 2B). The ensuing rigid-body motions in the TMDs result in a transition of the channel that alternately opens the lumen of the exporter to either the interior or exterior face of the cytoplasmic membrane. This is the basis of the alternate access model, in which the transporter switches between the "open" and "closed" states (122, 138). However, recent data also suggest that some ABC transporters do not fit within a single mechanistic model for generation of alternating access (23, 46, 69, 88).

Structures have now been solved for several ABC transporters, and a growing body of biochemical data is available. However, there remain a number of critical unresolved questions concerning their precise mode of action (62), particularly the molecular events that occur during the transition of the channel between open and closed states and during the energization of the transporter. ABC transporters are responsible for shuttling a vast variety of substrates across the membrane, from small peptides and drugs to long chains of polysaccharide or protein. It is still unclear how these molecules with diverse physical properties interact with the transporter. These questions will be answered only by further investigation using both biochemical and structural methods.

## PHYLOGENY OF ABC TRANSPORTERS INVOLVED IN GLYCAN EXPORT

The similarities in NBD sequences mean that ABC transporters are readily identified from genome databases, and this has been exploited to examine the distribution of glyco-ABC transporters. Homologs were identified by key word and BLAST searches of the NCBI database, using known NBD proteins involved in oligo- and polysaccharide export. These included Wzt (O-PS), KpsT (group 2 CPS), TagH (teichoic acid), PglK (N-linked glycoprotein), and MsbA (LPS lipid A flippase) homologs. The genomic context of each of the NBD "hits" was examined to ensure that they were affiliated with authentic glycan biosynthesis loci. To do this, the open reading frames surrounding each of the identified NBD homologs were examined for the presence of diagnostic genes encoding glycosyltransferases or enzymes involved in sugar-nucleotide precursor biosynthesis. The NBDs of glyco-ABC transporters analyzed here are listed in Table 1. The conserved ATP-binding portion of each NBD was used to establish the phylogenetic relationships between them. MalK (the well-characterized NBD protein from the maltose transporter [37]) was included as the outgroup. NBD homologs were identified in both Gram-positive and Gram-negative bacteria, as well as in archaea, indicating the widespread involvement of ABC transporters in glycan export in prokaryotes. The phylogenetic analysis revealed seven distinct clusters of NBDs (Fig. 3), and similar phylogenetic tree topologies were observed using both distance and parsimony algorithms, providing support for the analysis. With the exception of NBD proteins found in group G, the glyco-ABC transporters consist of independent TMD and NBD polypeptides. An interesting, and potentially diagnostic, feature of the corresponding NBDs is that most contain a variant signature motif with the sequence YSSGM. In contrast, group G proteins are classical "half-transporters," functioning as homodimers with each monomer containing one TMD and one NBD, and these contain the conventional LSGGQ signature motif.

Groups A and B contain NBD homologs from both O-PS and glycoprotein biosynthesis systems. They include representatives from Gram-positive and Gram-negative bacteria, as

TABLE 1. NBD homologs from oligo- and polysaccharide biosynthesis systems

| Species and NBD group | Strain or serotype | NBD | GenBank accession no. | Locus tag | Length (amino acids) | Organism classification |
|---|---|---|---|---|---|---|
| Group A | | | | | | |
| *Aggregatibacter actinomycetemcomitans* | Serotype e | ORF11 | BAA82537 | | 398 | Gammaproteobacteria |
| *Aeromonas salmonicida* subsp. *salmonicida* | A449 | AbcA | YP_001141286 | | 438 | Gammaproteobacteria |
| *Aneurinibacillus thermoaerophilus* | DSM10155/G$^+$ | Wzt | AAS49125 | | 435 | Firmicutes |
| | L420-91$^T$ | Wzt | AAS55714 | | 408 | Firmicutes |
| *Aromatoleum aromaticum* | EbN1 | | YP_157893 | ebA1593 | 483 | Betaproteobacteria |
| *Arthrospira maxima* | CS-328 | | ZP_03274399 | AmaxDRAFT_3223 | 472 | Cyanobacteria |
| *Aurantimonas manganoxydans* | SI85-9A1 | | ZP_01226056 | SI859A1_02283 | 424 | Alphaproteobacteria |
| *Burkholderia ambifaria* | MC40-6 | | YP_001807479 | BamMC406_0767 | 406 | Betaproteobacteria |
| *Burkholderia cenocepacia* | J2315 | Wzt | YP_002232239 | BCAL3130 | 437 | Betaproteobacteria |
| *Burkholderia phymatum* | STM815 | | YP_001858534 | Bphy_2313 | 450 | Betaproteobacteria |
| *Burkholderia pseudomallei* | 668 | | YP_001060111 | BURPS668_3099 | 465 | Betaproteobacteria |
| *Clostridium botulinum* | B strain Eklund 17B | | YP_001887433 | CLL_A3247 | 406 | Firmicutes |
| *Enterococcus faecalis* | V583 | | AAO81914 | EF_2182 | 405 | Firmicutes |
| *Escherichia coli* | O8 | Wzt | BAA28325 | | 404 | Firmicutes |
| | O9a | Wzt | BAA28332 | | 431 | Gammaproteobacteria |
| | O99 | Wzt | ACV53836 | | 433 | Gammaproteobacteria |
| *Geobacillus stearothermophilus* | NRS2004/3a | Wzt | AAR99607 | | 409 | Firmicutes |
| *Geobacillus tepidamans* | GS5-97$^T$ | Wzt | ABM68319 | | 395 | Firmicutes |
| *Geobacter uraniireducens* | Rf4 | | YP_001232537 | Gura_3813 | 711 | Deltaproteobacteria |
| *Hyphomonas neptunium* | ATCC 15444 | | YP_759524 | HNE_0796 | 417 | Alphaproteobacteria |
| *Klebsiella pneumoniae* | O12 | Wzt | AAN06493 | | 440 | Gammaproteobacteria |
| *Lyngbya* sp. | PCC 8106 | | ZP_01619961 | L8106_24925 | 468 | Cyanobacteria |
| *Mesorhizobium* sp. | BNC1 | | YP_676227 | Meso_3694 | 429 | Alphaproteobacteria |
| *Methanosarcina barkeri* | Fusaro | | YP_303935 | Mbar_A0371 | 410 | *Euryarchaeota* |
| *Nitrosomonas europaea* | ATCC 19718 | Wzt | NP_840568 | NE0483 | 451 | Betaproteobacteria |
| *Nitrosomonas eutropha* | C71 | | YP_746405 | Neut_0152 | 437 | Betaproteobacteria |
| *Nitrosospira multiformis* | ATCC 25196 | | YP_413079 | Nmul_A2398 | 488 | Betaproteobacteria |
| *Pseudomonas aeruginosa* | PAO1 | Wzt | AAC16668 | PA5450 | 421 | Gammaproteobacteria |
| *Pseudomonas putida* | KT2440 | | NP_743935 | PP_1779 | 405 | Gammaproteobacteria |
| | W619 | | YP_001748261 | PputW619_1387 | 402 | Gammaproteobacteria |
| *Pseudomonas syringae* pv. Tomato | DC3000 | | NP_790909 | PSPTO_1075 | 454 | Gammaproteobacteria |
| *Roseiflexus* sp. | RS-1 | | YP_001278383 | RoseRS_4089 | 870 | *Chloroflexi* |
| *Serratia marcescens* | O4 | Wzt | AAC00182 | | 441 | Gammaproteobacteria |
| *Sodalis glossinidius* | Morsitans | | YP_454797 | SG1117 | 434 | Gammaproteobacteria |
| *Stenotrophomonas* sp. | SKA14 | | ZP_05135042 | SSKA14_2119 | 475 | Gammaproteobacteria |
| *Sulfurimonas denitrificans* | DSM 1251 | | YP_394240 | Suden_1731 | 410 | Epsilonproteobacteria |
| *Vibrio cholerae* | AM-19226 | | ZP_04961260 | A33_0258 | 392 | Gammaproteobacteria |
| | O37 | | AAM22592 | | 429 | Gammaproteobacteria |
| *Xanthomonas axonopodis* pv. Dieffenbachiae | LMG695 | | AAZ08621 | | 329 | Gammaproteobacteria |
| *Xanthomonas campestris* pv. Campestris | B100 | Wzt | AAK53481 | | 428 | Gammaproteobacteria |
| *Xanthomonas oryzae* pv. Oryzae | BXO8 | Wzt | ABI93186 | | 410 | Gammaproteobacteria |
| | MAFF 311018 | | YP_449745 | XOO_0716 | 437 | Gammaproteobacteria |
| *Yersinia frederiksenii* | ATCC 33641 | | ZP_04633619 | yfred0001_9870 | 408 | Gammaproteobacteria |
| Group B | | | | | | |
| *Actinomyces odontolyticus* | ATCC 17982 | | ZP_02045164 | ACTODO_02054 | 430 | Actinobacteria |
| *Archaeoglobus fulgidus* | DSM 4304 | | NP_068882 | AF0041 | 237 | *Euryarchaeota* |
| *Cellulomonas flavigena* | DSM 20109 | | ZP_04366067 | CflaDRAFT_13740 | 411 | Actinobacteria |
| *Dokdonia donghaensis* | MED134 | | ZP_01051491 | MED134_13566 | 415 | Flavobacteria |
| *Flavobacterium johnsoniae* | UW101 | | YP_001192680 | Fjoh_0325 | 422 | Flavobacteria |
| *Mariprofundus ferrooxydans* | PV-1 | | ZP_01452528 | SPV1_07566 | 421 | Zetaproteobacteria |
| *Methanosaeta thermophila* | PT | | YP_843490 | Mthe_1066 | 399 | *Euryarchaeota* |
| *Methanosarcina acetivorans* | C2A | | NP_616118 | MA1177 | 409 | *Euryarchaeota* |
| *Methanosarcina mazei* | Gol | | NP_632680 | MM_0656 | 421 | *Euryarchaeota* |
| | Gol | | NP_634238 | MM_2214 | 504 | *Euryarchaeota* |
| *Microcoleus chthonoplastes* | PCC 7420 | | ZP_05026777 | MC7420_2165 | 433 | Cyanobacteria |
| *Rhizobium etli* | CE3 | Wzt | AAK51165 | | 443 | Alphaproteobacteria |
| | CIAT 652 | | YP_001976997 | RHECIAT_CH0000832 | 439 | Alphaproteobacteria |
| *Spirosoma linguale* | DSM 74 | | ZP_04492830 | SlinDRAFT_59260 | 419 | Bacteroidetes |
| *Streptococcus mutans* | UA159 | RgpD | NP_721238 | SMU.828 | 405 | Firmicutes |
| *Synechocystis* sp. | PCC 6803 | RfbB | NP_440220 | slr0982 | 430 | Cyanobacteria |
| Group C | | | | | | |
| *Erwinia tasmaniensis* | Et1/99 | RfbB | YP_001907278 | ETA_13390 | 246 | Gammaproteobacteria |
| *Escherichia coil* | SMS-3-5 | RfbB | YP_001744315 | EcSMS35_2267 | 246 | Gammaproteobacteria |
| *Klebsiella pneumoniae* | O1 | RfbB/Wzt | Q48476 | | 246 | Gammaproteobacteria |
| | O8 | RfbB/Wzt | Q48479 | | 246 | Gammaproteobacteria |
| *Rhodopseudomonas palustris* | BisB18 | | YP_534033 | RPC_4190 | 246 | Alphaproteobacteria |
| *Serratia marcescens* | O16 | Wzt | AAC98415 | | 246 | Gammaproteobacteria |
| Group D | | | | | | |
| *Aggregatibacter actinomycetemcomitans* | Serotype a | Orf4/Wzt | BAB03203 | | 250 | Gammaproteobacteria |
| | Serotype b | Orf11/Wzt | BAA19638 | | 245 | Gammaproteobacteria |

*Continued on following page*

TABLE 1—*Continued*

| Species and NBD group | Strain or serotype | NBD | GenBank accession no. | Locus tag | Length (amino acids) | Organism classification |
|---|---|---|---|---|---|---|
| | Serotype c | Wzt | BAA28136 | | 247 | Gammaproteobacteria |
| | Serotype f | Wzt | AAG49408 | | 248 | Gammaproteobacteria |
| *Actinobacillus pleuropneumoniae* | O1 | Wzt | AAR01223 | | 244 | Gammaproteobacteria |
| *Aeromonas hydrophila* subsp. *hydrophila* | ATCC 7966 | | YP_857403 | AHA_2899 | 252 | Gammaproteobacteria |
| *Agrobacterium radiobacter* | K84 | Wzt | YP_002546388 | Arad_4840 | 248 | Alphaproteobacteria |
| *Anaerocellum thermophilum* | DSM 6725 | | YP_002571988 | Athe_0053 | 244 | Firmicutes |
| *Anaerotruncus colihominis* | DSM 17241 | | ZP_02442862 | ANACOL_02160 | 239 | Firmicutes |
| *Azospirillum brasilense* | | Wzt | AAS83022 | pRhico005 | 260 | Alphaproteobacteria |
| *Bradyrhizobium* sp. | BTAi1 | | YP_001241520 | BBta_5660 | 249 | Alphaproteobacteria |
| *Brucella suis* | 1330 | RfbE | NP_697541 | BR0519 | 252 | Alphaproteobacteria |
| *Burkholderia cepacia* | AMMD | | YP_772638 | Bamb_0745 | 249 | Betaproteobacteria |
| *Caulobacter crescentus* | CB15 | | NP_419451 | CC_0634 | 250 | Alphaproteobacteria |
| *Clostridium methylpentosum* | DSM 5476 | | ZP_03705427 | CLOSTMETH_00138 | 244 | Firmicutes |
| *Coprococcus comes* | ATCC 27758 | | ZP_03799480 | COPCOM_01739 | 245 | Firmicutes |
| *Corynebacterium glutamicum* | ATCC 13032 | | YP_224499 | cg0248 | 263 | Actinobacteria |
| *Coxiella burnetii* | RSA 493 | RfbI | NP_819734 | CBU_0704 | 258 | Gammaproteobacteria |
| *Dichelobacter nodosus* | VCS1703A | | YP_001209256 | DNO_0336 | 233 | Gammaproteobacteria |
| *Dorea formicigenerans* | ATCC 27755 | | ZP_02234094 | DORFOR_00952 | 244 | Firmicutes |
| *Dorea longicatena* | DSM 13814 | | ZP_01995719 | DORLON_01714 | 252 | Firmicutes |
| *Erwinia chrysanthemi* | | Wzt | AAM33313 | | 243 | Gammaproteobacteria |
| *Escherichia coli* | O52 | Wzt | AAS99165 | | 324 | Gammaproteobacteria |
| *Escherichia hermannii* | YS-11 | Wzt | BAG06240 | | 248 | Gammaproteobacteria |
| *Eubacterium ventriosum* | ATCC 27560 | | ZP_02026283 | EUBVEN_01539 | 249 | Firmicutes |
| *Mycobacterium tuberculosis* | H37Rv | RfbE | NP_218298 | Rv3781 | 273 | Actinobacteria |
| *Pelotamaculum thermopropionicum* | SI | TagH | YP_001213125 | PTH_2575 | 266 | Firmicutes |
| *Pseudomonas fluorescens* | PfO-1 | | YP_349793 | Pfl01_4065 | 252 | Gammaproteobacteria |
| *Ruminococcus gnavus* | ATCC 29149 | | ZP_02042722 | RUMGNA_03526 | 244 | Firmicutes |
| *Ruminococcus torques* | ATCC 27756 | | ZP_01967310 | RUMTOR_00857 | 251 | Firmicutes |
| *Saccharopolyspora erythraea* | NRRL 2338 | | YP_001102478 | SACE_0200 | 272 | Actinobacteria |
| *Salmonella enterica* serovar Typhimurium | LT2 | | NP_459708 | STM0723 | 236 | Gammaproteobacteria |
| *Stenotrophomonas maltophilia* | R551-3 | | YP_002026879 | Smal_0491 | 241 | Gammaproteobacteria |
| *Vibrio cholerae* | O1 | RfbI | CAA42140 | | 250 | Gammaproteobacteria |
| *Xylella fastidiosa* | Dixon | | ZP_00651487 | XfasaDRAFT_1586 | 246 | Gammaproteobacteria |
| *Yersinia enterocolitica* | O:3 | RfbE | Q56903 | | 239 | Gammaproteobacteria |
| | O:5 | Wzt | AAT91802 | | 241 | Gammaproteobacteria |
| | O:9 | Wzt | CAE53859 | | 251 | Gammaproteobacteria |
| Group E | | | | | | |
| *Bacillus anthracis* | Ames | TagH | NP_847669 | BA5510 | 549 | Firmicutes |
| *Bacillus subtilis* | 168 | TagH | P42954 | BSU35700 | 527 | Firmicutes |
| | W23 | TarH | CAJ97406 | | 527 | Firmicutes |
| *Enterococcus faecalis* | V583 | TagH | Q831L8 | EF_2486 | 447 | Firmicutes |
| *Lactococcus lactis* | Il1403 | TagH | Q9CH26 | LL0915 | 466 | Firmicutes |
| *Leuconostoc citreum* | KM20 | TagH | YP_001728599 | LCK_01331 | 365 | Firmicutes |
| *Listeria monocytogenes* | EGD-e | | NP_464600 | lmo1075 | 333 | Firmicutes |
| *Staphylococcus aureus* subsp. *aureus* | COL | | YP_185576 | SACOL0694 | 264 | Firmicutes |
| Group F | | | | | | |
| *Actinobacillus pleuropneumoniae* | 5a | CpxA | AAB64445 | | 216 | Gammaproteobacteria |
| *Actinobacillus suis* | Strain SO4, serotype K1 | Wzt | AAO65487 | | 216 | Gammaproteobacteria |
| *Aeromonas hydrophila* | PPD134/91 | KpsT | AAM22566 | | 163[a] | Gammaproteobacteria |
| *Bordetella bronchiseptica* | RB50 | KpsT | NP_889466 | BB2930 | 245 | Betaproteobacteria |
| *Burkholderia pseudomallei* | K96243 | Wzt2 | YP_109399 | BPSL2804 | 218 | Betaproteobacteria |
| *Campylobacter jejuni* | NCTC 11168 | KpsT | YP_002344829 | Cj1447c | 220 | Epsilonproteobacteria |
| *Chlorobium phaeobacteroides* | DSM 266 | | YP_912756 | Cpha266_2344 | 217 | *Chlorobi* |
| *Citrobacter freundii* | | VexC | AAK14187 | | 213 | Gammaproteobacteria |
| *Escherichia coli* | K1 | KpsT | B42469 | | 218 | Gammaproteobacteria |
| | K5 | KpsT | P24586 | | 224 | Gammaproteobacteria |
| | K10 | KpsT | AAD31429 | | 216 | Gammaproteobacteria |
| | K54 | KpsT | AAC38079 | | 216 | Gammaproteobacteria |
| *Haemophilus influenzae* | Serotype b | BexA | CAA38734 | | 217 | Gammaproteobacteria |
| *Mannheimia haemolytica* | A1 | CpxA | AAF08240 | | 215 | Gammaproteobacteria |
| *Neisseria meningitidis* | Serogroup A strain Z2491 | CtrD | YP_002341738 | NMA0195 | 216 | Betaproteobacteria |
| | Serogroup B | CtrD | P32016 | NMB0074 | 216 | Betaproteobacteria |
| *Nitrococcus mobilis* | Nb-231 | | ZP_01127549 | NB231_02823 | 221 | Gammaproteobacteria |
| *Novosphingobium aromaticivorans* | DSM 12444 | | YP_496029 | Saro_0748 | 232 | Alphaproteobacteria |
| *Pasteurella multocida* | | CexA | AAF67272 | | 218 | Gammaproteobacteria |
| *Ralstonia metallidurans* | CH34 | | YP_587859 | Rmet_5731 | 221 | Betaproteobacteria |
| *Salmonella enterica* subsp. *enterica* serovar Typhi | | VexC/Wzt | BAA03198 | | 246 | Gammaproteobacteria |
| *Sinorhizobium meliloti* | 1021 | RkpS | CAC48975 | SM_b20832 | 219 | Alphaproteobacteria |
| *Yersinia intermedia* | ATCC 29909 | | ZP_04635682 | yinte0001_9460 | 221 | Gammaproteobacteria |

*Continued on following page*

TABLE 1—*Continued*

| Species and NBD group | Strain or serotype | NBD | GenBank accession no. | Locus tag | Length (amino acids) | Organism classification |
|---|---|---|---|---|---|---|
| Group G | | | | | | |
| *Arcobacter butzleri* | RM4018 | | YP_001489609 | | 566 | Epsilonproteobacteria |
| *Bordetella petrii* | DSM 12804 | MsbA2 | YP_001633454 | | 558 | Betaproteobacteria |
| | DSM 12804 | MsbA1 | YP_001631720 | | 615 | Betaproteobacteria |
| *Caulobacter crescentus* | CB15 | | NP_419124 | CC_0305 | 607 | Alphaproteobacteria |
| *Campylobacter coli* | RM2228 | WlaB | ZP_00367153 | CCO1205 | 566 | Epsilonproteobacteria |
| *Campylobacter jejuni* | NCTC 11168 | WlaB/PglK | YP_002344523 | Cj1130c | 564 | Epsilonproteobacteria |
| | NCTC 11168 | MsbA | YP_002344210 | Cj0803 | 580 | Epsilonproteobacteria |
| *Cyanothece* sp. | PCC 7425 | | YP_002482894 | Cyan7425_2172 | 610 | Cyanobacteria |
| *Escherichia coli* | | MsbA | CAA77839 | | 582 | Gammaproteobacteria |
| *Francisella novicida* | U112 | MsbA | YP_899220 | FTN_1606 | 609 | Gammaproteobacteria |
| *Helicobacter pylori* | J99 | Wzk | NP_223847 | | 578 | Epsilonproteobacteria |
| *Microcoleus chthonoplastes* | PCC 7420 | | ZP_05030053 | MC7420_5135 | 601 | Cyanobacteria |
| *Myxococcus xanthus* | DK 1622 | | YP_632878 | MXAN_4716 | 594 | Deltaproteobacteria |
| *Nostoc* sp. | PCC 7120 | HepA | AAC32400 | | 532 | Cyanobacteria |
| *Pseudomonas aeruginosa* | PAO1 | MsbA | AAG08382 | PA4997 | 603 | Gammaproteobacteria |
| *Wolinella succinogenes* | DSMZ 1740 | Wlab | CAE09222 | WS0052 | 563 | Epsilonproteobacteria |
| Others | | | | | | |
| *Bordetella parapertussis* | Bpp5 | WbmM | ABF72476 | | 393 | Betaproteobacteria |
| | Bpp5 | WbmN | ABF72475 | | 388 | Betaproteobacteria |
| *Myxococcus xanthus* | | RfbB | Q50863 | | 437 | Deltaproteobacteria |

[a] The start site was reannotated.

well as some archaea, and are unified by the presence of an extended C-terminal domain on the NBD. The sequences of the C-terminal domains suggest that they play diverse roles in cellular physiology. Here we use the term NBD to include ABC proteins that contain only the NBD domain as well as those that include additional functional (or putative functional) domains. Group A is larger than group B and contains mostly NBD homologs from low-G+C Gram-positive bacteria and Gram-negative proteobacteria. Group B is more diverse, containing examples from both high- and low-G+C Gram-positive bacteria as well as from a range of Gram-negative bacteria and archaea. Groups C, D, and F contain known and putative NBD homologs lacking an extended C terminus. Most examples in groups C and D are from O-PS assembly systems, but a few representatives from Gram-positive bacteria are also present. As with groups A and B, the significance of these NBDs clustering into two groups is unclear. Group E represents a relatively tight cluster of NBDs involved in the export of the polyol phosphate teichoic acids in Gram-positive bacteria. While these molecules are not oligo- or polysaccharides *per se*, their mode of synthesis is quite similar to that of many other bacterial glycoconjugates, and this justifies their inclusion here. Group F includes known and putative NBD homologs involved in the export of the "group 2" CPSs in Gram-negative bacteria. Finally, group G contains homologs of PglK and MsbA. This is the only group of half-transporters identified by our survey. MsbA proteins predominate due to the large number of known and highly conserved homologs in those bacteria that require LPS for viability. PglK is involved in the export of the general N-linked protein glycan of *C. jejuni*, and recent data have identified homologs from O-PS assembly systems in *Helicobacter pylori* (see below).

In most cases, information on NBD structure was available for representatives of each group to provide insight into the biosynthesis-export mechanism, as evaluated in the discussion below. Taking into consideration the available information, we have assigned prototype systems to each of the seven NBD groups (Table 2).

## ABC TRANSPORTERS WITH NBDs CONTAINING A CARBOHYDRATE-BINDING MODULE (CBM)—THE *E. COLI* POLYMANNOSE O-PS PARADIGM

ABC transporters involved in O-PS biosynthesis are comprised of separate TMD and NBD polypeptides, namely, Wzm and Wzt, respectively. Recent research with the *E. coli* O8 and O9a O-PSs has illustrated a pivotal role played by the ABC transporter (specifically the C-terminal extension on the NBD) in coordinating the biosynthesis and export phases of the assembly process (26, 31, 32). These NBDs are found in group A (Fig. 3).

The O-PSs of *E. coli* serotypes O8, O9, and O9a represent well-established examples of ABC transporter-dependent O-PS assembly systems. Identical O-PS structures are found in *Klebsiella pneumoniae* O serotypes, reflecting horizontal transfer of the corresponding loci (153). These O-PSs are homopolymers of mannose and differ in the linkage sequence that defines the polysaccharide repeat unit (Fig. 4). The gene clusters responsible for O-PS production in each serotype have a common organization and share many components, reflecting highly conserved (sometimes identical) steps in the biosynthesis of the polymannose O-PSs (70, 152, 154). Biosynthesis of the polymannose O-PS requires the activated precursor GDP-mannose and the activities of three dedicated mannosyltransferase enzymes, WbdA, WbdB, and WbdC (70). WbdB and WbdC are identical in each serotype, while WbdA varies significantly. The acceptor for mannosyltransferase activity is Und-PP–GlcNAc produced by the WecA enzyme, a GlcNAc: Und-P GlcNAc-1-phosphate transferase (61, 70, 126). WecA is involved in the biosynthesis of various O-PSs and enterobacterial common antigen, which, as the name suggests, is present in virtually all *Enterobacteriaceae*. There remains some debate about the precise order of action of the three mannosyltransferases and which enzymes contribute to the serospecific repeat unit domain (70, 163), but it is clear that WbdA, WbdB, and WbdC are collectively essential and sufficient for biosynthesis of the polymannose structures.

The polymannose O-PSs terminate with unique residues

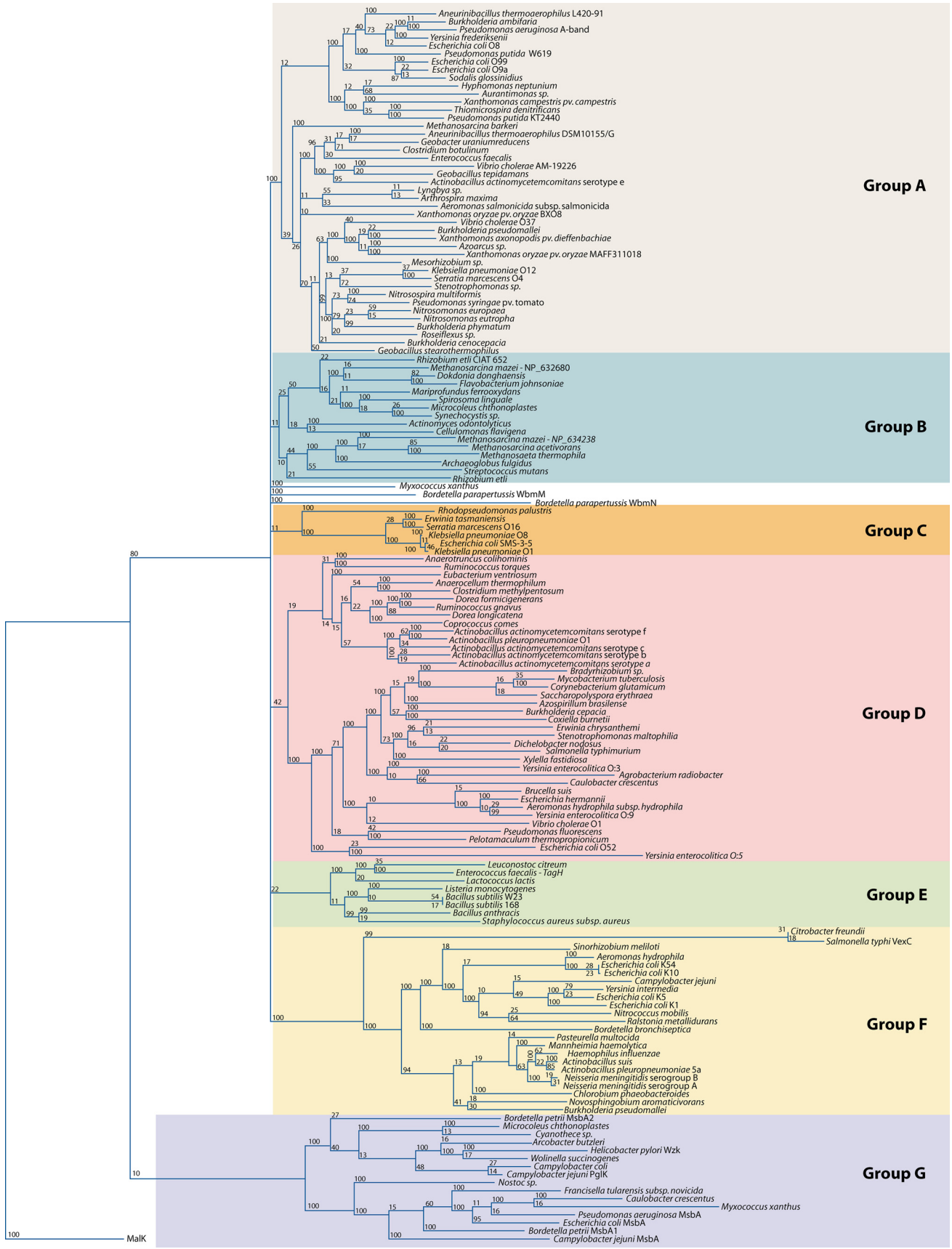


FIG. 3. Phylogenetic relationships among glyco-ABC transporters. Neighbor-joining analysis was carried out on 1,000 bootstrapped data sets, using the PHYLIP package (42). Trees were viewed using SplitsTree (59). Branch lengths represent relative distances, and bootstrap values are indicated. The tree is based on data available in the NCBI databases in January 2010.

TABLE 2. Representative ABC transporters from each of the seven phylogenetic groups

| Group | Substrates | Prototype | Structure of glycan[a] | Reference |
|---|---|---|---|---|
| A | O-PSs, glycoprotein O-linked glycans | *Escherichia coli* O9a Wzt protein (O-PS) | Me-P-? [→**2)-α-Man-(1→2)-α-Man-(1→3)-α-Man-(1→3)-α-Man-(1→**]$_n$→3)-α-Man-(1→3)-α-Man-(1→3)-β-GlcNAc | 26, 163 |
| B | O-PSs, CPSs?, glycoprotein glycans | *Rhizobium etli* CE3 Wzt protein (O-PS) | α-2,3,4-Tri-*O*-Me-Fuc-(1→[→4)-α-GlcA-(1→4)-α-Fuc-(1→]$_n$→3)-α-Fuc-(1→3)-β-Man-(1→3)-β-QuiNAc<br>3<br>↑<br>α-3-*O*-Me-Tal-1 | 45 |
| C | O-PSs, others | *Klebsiella pneumoniae* O2a Wzt protein (O-PS) | [→**3)-β-Gal*f*-(1→3)-α-Gal-(1→**]$_n$→3)-β-Gal*f*-(1→3)-α-Gal-(1→3)-β-GlcNAc (D-galactan I) | 73, 74, 174, 175 |
| D | O-PSs, others | *Yersinia enterocolitica* O:3 Wzt protein (O-PS) | [→**2)-β-Alt-(1→**]$_n$ | 56 |
| E | Teichoic acids | *Bacillus subtilis* 168 TagH protein (wall teichoic acid) | [GroP]$_n$→[GroP]$_{1-3}$→4-ManNAc-(1→4)-β-GlcNAc<br>↑<br>*X (X = H, D-Ala, α-Glc) | 18 |
| F | "Group 2"-related CPSs | *Escherichia coli* K1 KpsT protein | [→8)-α-Neu5Ac-(2→]<br>7/9<br>↑<br>**O*-Ac | 93 |
| G | Lipid A, glycoprotein N-linked glycans | *Campylobacter jejuni* (PglK) | α-GalNAc-(1→4)-α-GalNAc-(1→4)-α-GalNAc-(1→4)-α-GalNAc-(1→4)-α-GalNAc-(1→3)-α-Bac2,4diNAc<br>3<br>↑<br>β-Glc-1 | 178 |

[a] Where the structure of the substrate includes a repeat unit domain, the relevant region is shown in bold. All sugars are pyranoses in the D configuration, unless otherwise indicated. Ac, acetyl; Ala, alanyl; Alt, 6-deoxy-L-altrose; Bac2,4-diNAc, 2,4-diacetamido-2,4,6-trideoxyglucose (bacillosamine); Fuc, fucose; Gal, galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcA, glucuronic acid; GlcNAc, *N*-acetylglucosamine; Gro, glycerol; Man, mannose; ManNAc, N-acetylmannosamine; Neu5Ac, *N*-acetylneuraminic acid (sialic acid); Me, methyl; P, phosphate; QuinNAc, 2-*N*-acetamido-2,6-dideoxyglucose (*N*-acetylquinovosamine); Tal, 6-deoxytalose. *, variable (nonstoichiometric) substituents.

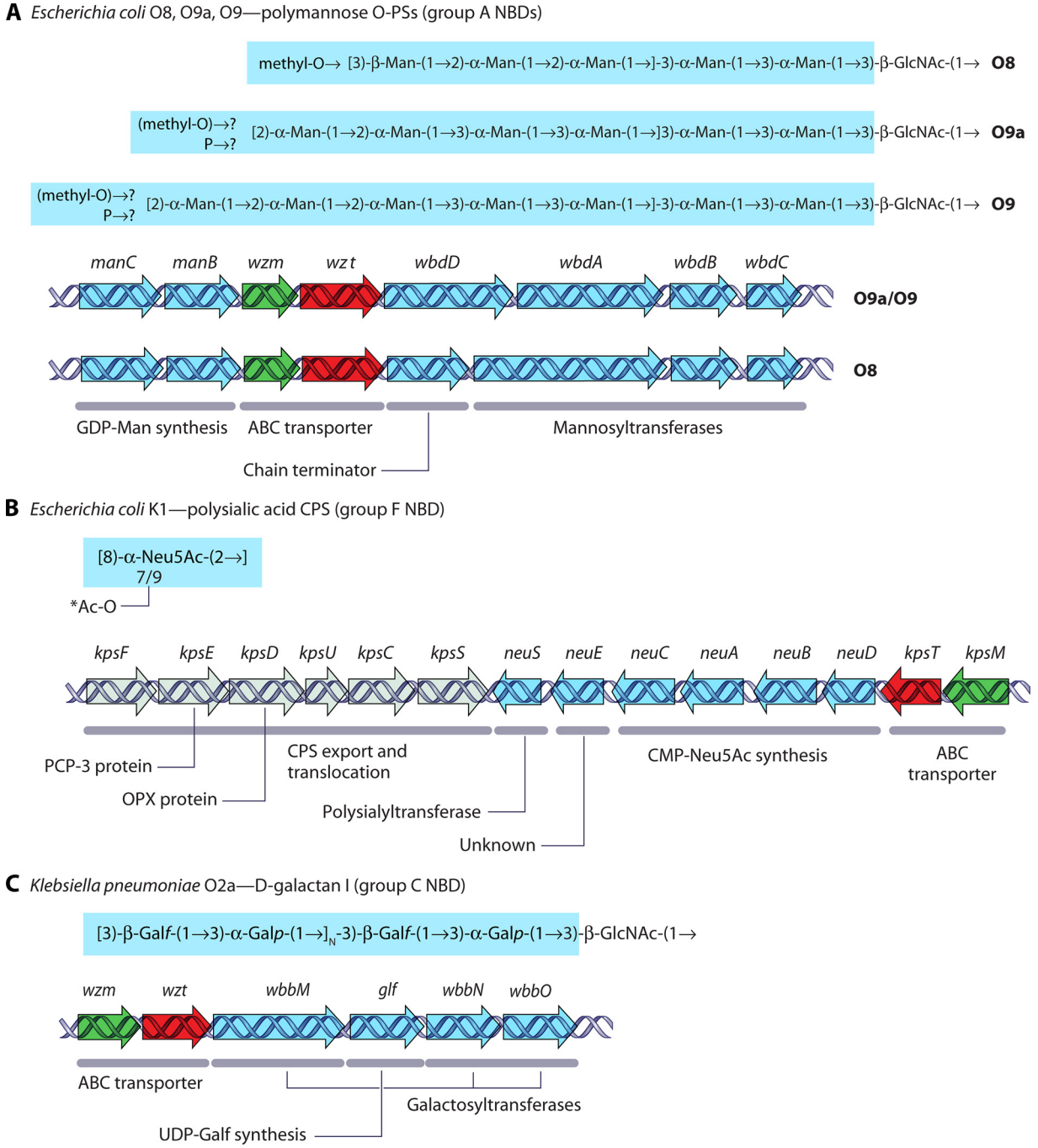


FIG. 4. Schematics of genetic loci and gene functions for well-characterized examples of glycan assembly systems that feature an ABC transporter. For each ABC transporter, the gene encoding the NBD protein is shown in red and the corresponding gene for the TMD is shown in green. Genes encoding glycan biosynthesis proteins are highlighted in blue, and the products of their activities are shaded in the structures provided in each panel. (A) Group A NBDs. (B) Group F NBDs. (C) Group C NBDs.

(i.e., ones not found within the repeat units) on the nonreducing termini of the polysaccharide chains, and these are added by chain-terminating WbdD enzymes (26). A combination of structural and biochemical data identified these terminating residues as a methyl group (*E. coli* O8) or a phosphate plus a methyl group in an unknown organization (serotypes O9 and O9a) (26, 86, 163). In serotype O8, WbdD is a methyltransferase, whereas the O9/O9a WbdD homolog is a bifunctional kinase-methyltransferase, with the kinase activity being a prerequisite for the addition of the methyl residues (26). Most O-PSs exhibit a range of chain lengths (in what is often called a "modal" distribution), which can vary quite considerably and give rise to the characteristic ladder patterns seen when LPS extracts are separated by SDS-PAGE. The chain-terminating

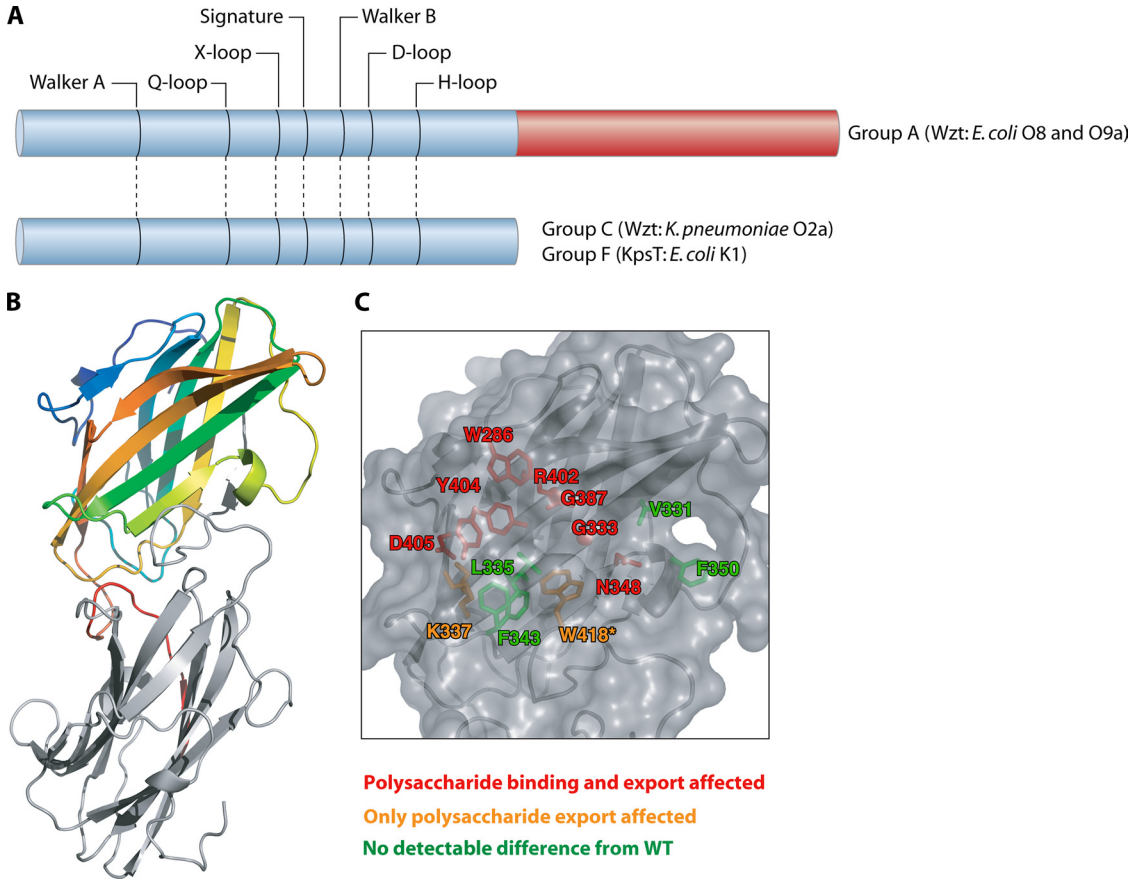


FIG. 5. Structure and organization of NBDs from glyco-ABC transporters. (A) Three representative NBDs all contain the highly conserved protein sequence motifs that are essential for ABC transporter function. In *E. coli* O8 and O9a (group A), the NBD contains an additional C-terminal extension (highlighted in red) that contains a CBM. Two characterized representative NBDs that lack extended C termini are shown for comparison. (B) The crystal structure of the C-terminal domain of the O9a NBD reveals a dimer in which each monomer is a modified Ig fold typical of certain CBM families. (C) The CBM contains a pocket essential for O-PS binding that is conserved in closely related homologs. The transport phenotypes of site-directed mutants identify critical roles for aromatic residues, suggesting that the CBM binds its substrate through a ring-stacking mechanism. (Panels B and C were reprinted from reference 31.)

activity of WbdD is central to the establishment of the relatively limited modal chain-length range seen in the polymannose O-PSs. Overexpression of WbdD enhances chain termination frequency and reduces O-PS chain length in serotypes O8 and O9a, suggesting that the stoichiometry of the various biosynthesis components and their potential interactions may be important in determining O-PS chain length (26). Critical to this process is a physical interaction between WbdA and the C-terminal (noncatalytic) region of WbdD, which targets an otherwise soluble WbdA mannosyltransferase to the membrane, where it must reside for O-PS biosynthesis to occur (27). WbdD-mediated chain termination is also essential for coupling biosynthesis to export. *E. coli* O9a *wbdD* mutants defective in kinase-methyltransferase activities can synthesize the O9a polysaccharide but are unable to export it (26). This phenotype is essentially identical to that of an ABC transporter (Δ*wzm* Δ*wzt*) mutant. In both cases, the accumulation of polymer in the absence of export is toxic to the cell, and conditional (GDP-mannose synthesis) mutations are required to avoid second-site mutations that relieve the stress (26, 33).

A key element in understanding the function of the polymannose O-PS ABC transporters came from the observation that the O-PS O8 and O9a transporters are serotype specific. Although the Wzm proteins can be exchanged, albeit with reduced export efficiencies, the Wzt components cannot be exchanged between serotypes (33). The specificity observed in substrate recognition and transport lies within the Wzt components. Comparison of the serotype O8 and O9/O9a Wzt protein sequences and functional data from various chimeric proteins have established that the Wzt proteins contain two functional domains (33). The conserved (and exchangeable) N-terminal domain encompasses the characteristic motifs that define NBDs, whereas the variable C-terminal region confers serospecificity (Fig. 5). Both domains are required for export, but surprisingly, a functional transporter exists when the two domains are expressed separately in chromosomal Δ*wzt* mutants (31). The C-terminal domain provides the essential connection between polymer chain termination and export. Biochemical experiments established that the C-terminal domain of Wzt binds O-PS and, furthermore, that glycan binding requires the chain-terminating residues added by WbdD. The structure of the C-terminal domain of Wzt from serotype O9a

was determined by X-ray crystallography (Fig. 5B), revealing a dimer in which the interface between monomers is stabilized by the exchange of one β strand (B9) and the continuation of a single hydrophobic core through both monomers (31). The structure is a β sandwich with an immunoglobulin (Ig)-like topology, a motif seen in a number of carbohydrate-binding modules (CBMs), which serve to bring catalytic domains of carbohydrate-active proteins into close proximity with their cognate substrates. Based on sequence similarity, CBMs are currently divided into more than 50 families, and the number is growing (www.cazy.org/). The β-sandwich fold is the most common fold seen in CBMs, and at least 10 of these families (at current count) have an Ig-like topology (54). The C-terminal domain of Wzt provides a broad substrate-binding face where aromatic residues play a critical role (Fig. 5C), potentially by creating opportunities for ring-stacking interactions (31). Despite a common function shared by the CBMs of *E. coli* O8 and O9a, their sequences are quite different (21% identity and 43% similarity).

It is important that there is currently no direct evidence that the O-PS actually passes through the TMD of the ABC transporter. Although this is the most plausible explanation for the process and is consistent with relevant mutant phenotypes, it is still possible that the ABC transporter participates indirectly in exporting the polymer. Regardless, two critical unresolved questions exist. The first is how the Und-PP-linked intermediate traverses the inner membrane. The lipid-linked O-PS intermediate is a complex substrate with both hydrophobic and polar regions. It is conceivable that the polar and hydrophobic domains are handled differently. For example, the polar chain could pass through the hydrophilic channel, while the lipid acceptor may remain confined to the adjacent membrane core. A similar solution was proposed to address the same issue in the export of lipid A via MsbA (125). These are difficult questions to test experimentally without a structural snapshot that captures export in mid-progress. The second unresolved question is precisely how the C-terminal CBM of Wzt functions overall in the export process for the polymannose O-PSs. We currently envisage two possibilities. In one model, the binding domain introduces the polymer into the transport channel. However, this scenario is complicated by comparison to known ABC transporters, which predict that the C-terminal binding domain of the NBD is separated from the TMD by a considerable distance (~25 Å in the case of $MalFGK_2$ [35]). Establishing the position of the CBM in the context of the whole ABC transporter would provide important insight, but attempts to crystallize the full-length Wzt protein have proved unsuccessful. However, it is also possible that the CBM does not lead the substrate to the export channel. In an alternative model, the CBM may be required to disengage the nascent O-PS from the assembly (glycosyltransferase) enzymes to allow it to enter the export pathway. The recognition of nonreducing terminal residues would ensure that only terminated chains that are competent for export are released. Either scenario is consistent with the phenotypes of cells lacking the chain-terminating methyltransferase-kinase. It is unknown whether the modified nonreducing terminus is the first part of the polymer to enter the export channel or the last. In the latter scenario, export could begin before synthesis is complete. Synthesis and export can be uncoupled temporally in a conditional experimental situation, i.e., polymer completed in the absence of the ABC transporter can be exported once expression of the transporter is activated (77). This would be the equivalent of post-translational export of secreted proteins. However, the coordination suggests that tight coupling (akin to that of cotranslational export) is more likely to be the true physiological situation.

## OTHER SYSTEMS WITH NBDs CONTAINING PUTATIVE CBMs

Several other group A NBDs come from assembly systems for which there is biochemical evidence supporting an assembly process that follows the *E. coli* polymannose O-PS model. The *E. coli* O99 and *Pseudomonas aeruginosa* A-band O-PSs can be considered variants of *E. coli* O8/O9a O-PSs in terms of their biosynthesis. Both antigens have a backbone of poly-D-rhamnose (not the more usual L-rhamnose), and in O99, this is modified with glucose residues (9, 113). The activated precursor for the polyrhamnose backbone is GDP-D-rhamnose, which is synthesized from GDP-D-mannose in a two-step process involving a dehydratase and a reductase (71). The serotype O99 rhamnosyltransferases share significant levels of similarity with the corresponding mannosyltransferases from *E. coli* O8/O9a (113). In the structure of the O99 antigen, novel nonreducing terminating residues were not reported. However, the gene cluster encodes a predicted protein (WejH) sharing high levels of similarity (44% identity and 62% similarity) with the WbdD protein from *E. coli* O9a. The critical methyltransferase and kinase motifs are conserved, suggesting that chain termination follows a model very similar to the one reported for *E. coli* O9a. The A-band glycan contains methyl substitutions, and available data suggest that the methyl groups may be present at the nonreducing terminus (9). Genes encoding two putative methyltransferases were identified downstream and are transcribed in the opposite orientation from that of Wzt in *P. aeruginosa* (71). The A band is a conserved LPS-linked antigen and is typically coexpressed with a serospecific O-PS that is formed by a Wzx/Wzy-dependent pathway (71). Rhamnose O-PSs are common in plant-pathogenic species of *Pseudomonas*, and in at least some cases, nonreducing terminal methyl groups have been identified (181). The NBDs from two strains of *Pseudomonas putida* were included in the analysis, and they grouped with the Wzt homologs from *E. coli* O8, O9a, and O99 (group A).

It is not clear whether the O99 genetic locus also encodes the glucosyltransferases that modify the main chain. It would be particularly interesting to know if these residues are added before or after export across the cytoplasmic membrane. For the Wzx/Wzy-dependent O-PSs of *Shigella* and *Salmonella*, modifications, including glucosylation and O-acetylation, are encoded by serotype-converting lysogenic phages (3). The glucosylation occurs at the periplasmic face of the membrane while the nascent O-PS remains attached to its Und-PP carrier. The donor for glucosylation is Und-P-glucose, which is exported by a dedicated "flippase" protein via an unknown mechanism (51, 76). We are not aware of any confirmed examples of phage seroconversion involving an O-PS assembled via an ABC transporter-dependent

mechanism, but there are no features of the pathway that would necessarily preclude such a process.

An obvious issue arising from the examples above is whether all O-PSs assembled by this type of system (and having a group A NBD) terminate with nonreducing terminal methyl groups. The answer is no. The *K. pneumoniae* O12 O-PS has a repeat unit of rhamnose and GlcNAc, with a 3-deoxy-D-manno-octulosonic acid (Kdo) residue at its nonreducing terminus (163). It is likely that many more examples exist, but polysaccharide structural data tend to focus on the repeat unit, and the residues found only at the terminus of each chain can easily go undetected.

Another fundamentally important question is whether this type of system extends beyond O-PS biosynthesis. The answer is yes. Nonreducing terminal modifications analogous to those terminating the *E. coli* O8 and O9a O-PSs have also been identified on the S-layer glycans of a number of species of *Bacillaceae* (131). Others have already noted obvious similarities in the biosynthesis of some S-layer glycans and ABC transporter-dependent O-PSs (102, 134, 149). The ABC transporters from these S-layer glycan systems have separate TMD and NBD polypeptides resembling Wzm and Wzt, respectively (102). Furthermore, NBD proteins from well-investigated S-layer glycan assembly systems from *Geobacillus stearothermophilus* NRS2004/3a, *Geobacillus tepidamans* GS5-97$^T$ (180), and *Aneurinibacillus thermoaerophilus* strains L420-91$^T$ and DSM10155/G$^+$ are found in group A. The S-layer glycan from *G. stearothermophilus* strain NRS2004/3a contains an L-rhamnose trisaccharide repeat unit, and the chain of 15 repeat units is terminated with a 2-*O*-methyl group (133). The glycan chain is elongated on a Und-PP-galactose acceptor (149) and is linked to the protein via an O-glycosidic linkage (133). Interestingly, WsaE, the enzyme responsible for the *S*-adenosylmethionine-dependent methylation reaction, is a multifunctional protein that also has rhamnosyltransferase activity (150). The physical coupling of elongation and termination activities may be advantageous in regulating the glycan chain length. While biosynthesis information is limited for the other representatives, the structural theme is conserved: they possess a glycan capped with a nonreducing terminating residue. In *G. tepidamans* GS5-97$^T$, a chain of 20 repeat units containing D-fucose and L-rhamnose is O linked to two serine and threonine residues in the S-layer protein (63). The glycan is terminated by two substitutions, i.e., *N*-acetylglucosamine and *N*-acetylmuramic acid. The latter sugar is normally associated only with peptidoglycan. The *A. thermoaerophilus* L420-91$^T$ S-layer glycan contains an O-linked (to threonine) glycan with a complicated hexasaccharide repeat unit containing a backbone of D-rhamnose with D-*N*-acetylfucosamine side chains. It is terminated with a 3-*O*-methyl group (132). All evidence points to a conserved strategy for coupling biosynthesis and export in these systems and in the related O-PSs, such as those from *E. coli* O8/O9a. It would be interesting to determine whether the extended C-terminal domains on the corresponding S-layer glycan Wzt proteins contain cognate CBMs and thus warrant further investigation.

## VARIATIONS ON THE THEME—DIVERSITY IN FUNCTION OF EXTENDED C-TERMINAL DOMAINS IN GROUP B NBDs?

The O-PS of the plant symbiont *Rhizobium etli* CE3 contains multiple methyl groups, including three at the nonreducing terminus (45, 84). The extent of methylation varies according to environmental conditions, specifically, the presence of the host plant or seed exudate, and the internal methyl groups facilitate symbiosis. Multiple methyltransferase domains have been identified in gene products from the biosynthesis cluster (85, 104). The Wzt homolog from *R. etli* CE3 is found in group B and contains an extended C terminus, suggesting, at first glance, that this system follows the model established for O-PS and S-layer glycans described above. However, unlike the situation with *E. coli* O8/O9a O-PS biosynthesis, chain-length regulation and O-PS export are not dependent on nonreducing terminal methylation in *R. etli* CE3 (104). The role of the *R. etli* CE3 NBD C-terminal domain must therefore differ. The primary sequence of the C-terminal domain offers no clues to its function; from sequence data alone, it is not possible to rule in, or out, a novel CBM participating in a different means of assembly. Interestingly, *R. etli* provides an example of O-PSs possessing a homogenous chain length, with 5 repeat units (45), rather than the more typical modal range of chain lengths. It would be interesting to know whether the extended C-terminal domain of the NBD plays a role in establishing this homogeneity.

In some cases, bioinformatic analyses identified additional interesting putative functions for the C-terminal domains of group B NBDs. In three cases, NBD proteins, from *Geobacter uraniireducens*, *Roseiflexus* sp., and *Methanosarcina mazei*, were fused to C-terminal domains containing catalytic motifs. The additional motif in *Geobacter uraniireducens* is a putative sulfotransferase sharing some similarity (15% identity and 46% similarity) with NodH from *Rhizobium tropici* (82). NodH is a sulfotransferase (16) involved in modification of the lipochitooligosaccharide nodulation factors in *Rhizobium* sp. (40, 137). In *Roseiflexus* sp., the NBD homolog is fused to a predicted methyltransferase domain sharing similarity with RfbT from *Vibrio cholerae* O1 (151). RfbT is responsible for addition of the serogroup-specific methyl group to the O-PS of *V. cholerae* O1 (see below). Both NBD homologs contain an extended C-terminal region between the core NBD domain and the predicted catalytic domain. The NBD homolog from *Methanosarcina mazei* (MM_2214) is also fused at its C terminus to a predicted catalytic domain, in this case an acetyltransferase. It is tempting to speculate that the fused (putative) enzymatic domains found on the *G. uraniireducens*, *Roseiflexus* sp., and *M. mazei* NBD proteins may be involved in the addition of nonreducing terminal modifications to the polysaccharides synthesized. It would be interesting if, in these examples, different domains of the same polypeptide are responsible for both modification of the polysaccharide and its subsequent recognition of the modification prior to polysaccharide export. The biosynthetic consequences and possible advantages of such an arrangement are unclear.

A number of predicted NBDs found in group B are from the *Euryarchaeota*, including members of the genera *Methanosarcina*, *Archaeoglobus*, and *Methanosaeta*, and they appear to be

involved in protein glycosylation. In all cases, the genetic loci contain open reading frames encoding predicted homologs of an STT3 oligosaccharyltransferase homologous to PglB from *C. jejuni* (21). No information is available concerning the structure of the resulting glycans, so interpretation of the significance and roles of the extended C-terminal domains on the corresponding NBDs is currently not possible.

In *Aeromonas salmonicida*, the group B NBD appears to provide a direct connection between O-PS and S-layer glycan biosynthesis and illustrates a different activity for the NBD C-terminal extension. This bacterium produces homogenous (nonmodal) O-PS chain lengths (22), and the NBD protein (AbcA) is required for O-PS export (24). AbcA also contains an extended C-terminal domain that contains a putative leucine zipper, which is not essential for O-PS export but is required for proper transcription of the surface layer protein gene *vapA* (25, 100). Conversely, the activity of the core NBD portion of AbcA is confined to O-PS export. In this organism, the S layer is coupled to the cell surface via the O-PS chains (101), so the coupling of O-PS and S-layer protein expression may play an important physiological role. These NBDs present fascinating subjects for further analysis.

Three group B NBD proteins are particularly intriguing in terms of the current models for O-PS and CPS biosynthesis and assembly. The NBD homologs from *Dokdonia donghaensis*, *Mariprofundus ferrooxydans*, and *Synechocystis* sp. cluster quite closely. In each case, the NBD-encoding gene is located downstream of predicted open reading frames encoding OPX (outer membrane polysaccharide export) and PCP-2a homologs. PCP-2a and OPX proteins are inner and outer membrane components, respectively, of a transenvelope assembly-export machinery for Wzx/Wzy-dependent *E. coli* CPS (29, 32, 173). Again, it is difficult to interpret these findings in the absence of information concerning the glycoconjugates from these loci. Perhaps these organisms represent completely new twists on existing glycan assembly pathways.

## EXPORT OF GLYCANS VIA AN ABC EXPORTER LACKING A C-TERMINAL EXTENSION ON THE NBD PROTEIN—ASSEMBLY OF GROUP 2 CPSs

Group F contains known and putative NBD homologs involved in "group 2" CPS export in *E. coli*. They lack an extended C-terminal domain (Fig. 5). The group 2 designation identifies several CPSs of *E. coli* whose export involves an ABC transporter; "group 1" CPSs use a Wzx/Wzy-dependent process (173). Most work has been done with the *E. coli* K1 and K5 CPS assembly systems. In fact, *E. coli* K1 carries the first polysaccharide biosynthesis cluster to be cloned (141). The "group 2" name is often extended to CPS assembly systems in other bacteria which share the same spectrum of functional components. Examples include *Neisseria meningitidis*, *H. influenzae*, and *C. jejuni*.

In the *E. coli* paradigm, KpsM and KpsT form the TMD and NBD of the ABC transporter, respectively, and are required for the transport of the CPS across the inner membrane (80, 81, 108, 109, 116). Initial evidence concerning the role of KpsMT in CPS export came from the observation that mutants in either *kpsM* or *kpsT* accumulated CPS polymers within the cytoplasm (108, 109, 116). All of the CPS NBDs assigned to group F lack C-terminal extensions, and the corresponding ABC transporters show no apparent specificity for the repeat unit structures of their glycan substrates. Examples have been reported where the export apparatus is functionally interchangeable between different CPS serotypes, or even between different bacterial species (10, 87, 112, 128, 129, 169). Without molecular recognition, how can coupled polymerization and export be achieved? One possibility is that a presently unknown conserved part of the substrate might be recognized as an export signal. The diacylglycerol phosphate moieties found at the reducing terminus in this class of CPSs represent strong candidates if a molecular signal is indeed required (15). However, an attractive recent model invokes a system where there is no specific polymer recognition event in the coupling of synthesis to export. Instead, it is proposed that the system is coordinated by direct heterotypic protein-protein interactions, obviating the need for any conserved feature in the CPS export substrate (146, 162). Several studies have established the existence of multiprotein complexes involving proteins required for biosynthesis and export (94, 127, 146). In the case of *E. coli* K1, CPS biosynthesis occurs within a complex that effectively shields the nascent polymer from the cytoplasmic milieu (146). The multiprotein complex provides the structural context for coupling polymerization and export.

Synthesis of the *E. coli* K1 CPS occurs on the cytoplasmic face of the inner membrane, and several dedicated proteins participate in the assembly pathway (Fig. 4). The mechanism of initiation of CPS synthesis is also not yet resolved, but four gene products, NeuE, NeuS, KpsC, and KpsS, are required for *de novo* synthesis of cytoplasmic K1 CPS in *E. coli* (5). The polysialyltransferase (NeuS) is responsible for polymer extension through sequential glycosyl transfers to the nonreducing terminus of the growing glycan (147, 148, 164). The resulting CPSs contain a reducing terminal diacylglycerol. This could be attached through a Kdo linker residue (43), although a reducing terminal Kdo has not been identified on all of the related CPSs (49). It is not known at which point in CPS synthesis the polymer is transferred to the lipid anchor. For this system, the involvement of Und-linked intermediates is still controversial, and it is possible that diacylglycerol phosphate serves as the acceptor for chain initiation and extension (173). The KpsC protein has been implicated in the proper export of the CPS to the outer membrane. It has been proposed at various times that KpsC may be involved in Kdo addition and/or attachment of the lipid moiety to the reducing terminus. However, an *N. meningitidis* serotype B (polysialic acid) mutant deficient in its KpsC homolog can polymerize lipidated CPS in the absence of export (161). The recent proposal (162) that KpsC provides a pivotal adaptor between the biosynthetic enzymes and the ABC transporter seems a plausible alternative function. Two-hybrid data established that KpsC oligomerizes and interacts with both NeuS and KpsE (PCP-3 protein) (146). KpsC proteins are well distributed in bacteria with "group 2-like" CPSs (162).

In ABC transporter-dependent CPS biosynthesis, the final steps of translocation of the polymer across the periplasm and outer membrane require two dedicated proteins (8, 17, 94, 111, 140). In *E. coli*, these proteins include KpsE (a PCP-3 protein) and KpsD (an OPX representative). In the working model, these proteins are proposed to come together to form a struc-

ture analogous to a tripartite drug efflux pump (32). Interestingly, the clustering of NBD homologs in group F shows remarkable similarity to the phylogenetic relationships of the corresponding PCP-3 proteins (32), suggesting that important molecular recognition elements may be required for functional interactions between this class of ABC transporters and their cognate PCP and OPX (outer membrane channel) partners (142). We examined KpsC proteins and determined that they do not follow the same discrete phylogenetic clustering observed for NBD, PCP, and OPX proteins from the same organisms. While this certainly does not rule out a role for KpsC proteins as adaptors linking biosynthesis and export, it does not offer direct support for a cognate pairing between a given KpsC protein and the ABC transporter.

## O-PS EXPORT USING GROUP C AND D NBDs LACKING AN EXTENDED C TERMINUS—IS IT SIMILAR TO THE CPS SYSTEM?

Groups C and D contain NBDs from a variety of different O-PS biosynthesis systems, and they all lack a C-terminal extension (see the group C representative in Fig. 5A), suggesting that the mode of coupling polymerization and export is fundamentally different from the *E. coli* polymannose O-PS paradigm. Biochemical data are available for one example, the D-galactan I O-PS found in several O serotypes of *K. pneumoniae*. D-Galactan I is a heteropolysaccharide containing the pyranose and furanose forms of galactose, arranged in a disaccharide repeat unit (73, 74, 174) (Fig. 4). D-Galactan I defines the O2a antigen of *K. pneumoniae*, and in some isolates, it is the only O-PS. In other isolates, D-galactan I may also be extended (or capped) by an additional polysaccharide domain with a different structure, which can define a new serotype (e.g., O1 and O2c) (174, 175). Further diversification of serotypes can occur through additional modifications (e.g., serotype O8 is an O-acetylated version of serotype O1 [67]). O-PSs isolated from these strains contain a mixture of molecules carrying only D-galactan I and others comprising short chains of D-galactan I capped by the additional antigen (73, 74). The activated precursors for D-galactan I biosynthesis are UDP-galactopyranose and UDP-galactofuranose (75). Synthesis proceeds on a Und-PP–GlcNAc acceptor formed by WecA (28, 52, 163), and it requires three dedicated galactosyltransferases (WbbM, WbbN, and WbbO), whose precise contributions are only partially resolved. These steps essentially parallel the synthesis of the polymannose O-PSs, with the exception that there is no obvious terminating residue at the nonreducing terminus.

In the dendrogram in Fig. 3, NBDs from the *K. pneumoniae* D-galactan I systems are represented by the homologs from serotype O1 (whose sequence is identical to that of the NBD from O2a) and serotype O8 (which shows some sequence diversity) (67). *Serratia marcescens* O16 also contains D-galactan I (107), and the genetic locus is well conserved (156). The exception is the glycan synthesized by *Rhodopseudomonas palustris*, which likely contains L-rhamnose due to the presence of the *rml* (dTDP-L-rhamnose biosynthesis) genes immediately upstream of *wzm* and *wzt*. The lack of an extended C-terminal domain in the NBD of the D-galactan I O-PS exporter (Fig. 5A), together with the absence of chain-terminating modifications at the nonreducing terminus of the corresponding polymer, dictates a mechanism linking chain extension and export that differs from the *E. coli* O8/O9a paradigm. The model proposed for group 2 CPSs (see above) may be a better representation. Unlike the polymannose systems, where WbdD activity determines O-PS chain length, the chain length of D-galactan I is determined by the stoichiometry of the ABC transporter and the O-PS biosynthesis machinery (77). As a consequence, polymerization and export are obligatorily coupled in this system, analogous to cotranslational protein export. In the absence of the ABC transporter, aberrantly long, "unregulated" O-PS chains are synthesized, but these are not substrates for export if expression of the transporter is activated. It is unknown whether this is due only to the disconnection of synthesis and export or reflects a limitation in the maximum size of nascent polymers that the ABC transporter is able to export. Regardless, the equivalent of posttranslational protein export is not possible, unlike the case for polymannose O-PS systems. The modal distribution of O-PS chain lengths in LPS molecules containing D-galactan I is considerably more heterogeneous than that seen in O8 and O9a strains, and this may reflect the absence of a discrete chain-terminating enzymatic reaction. Interestingly, the disconnection of group 2 CPS synthesis can also affect polymer chain length in *Sinorhizobium meliloti* Rm1021 (98), so the phenotypes seen in D-galactan I O-PS mutants may reflect a broader phenomenon.

If the *E. coli* O8 and O9/O9a NBDs impart serospecificity by recognizing the nonreducing terminal modifications on the substrate, what does the *K. pneumoniae* O2a transport system recognize? Important insight comes from the observation that the *K. pneumoniae* O2a ABC transporter can complement a Δ*wzm* Δ*wzt* mutant of *E. coli* O9a to restore polymannose O-PS export (77). This led to the prediction that the *K. pneumoniae* O2a ABC transporter actually exports a wide variety of O-PS structures from different species, independent of repeat unit structures, like the case for the CPS ABC transporters containing group F NBDs. How is this achieved, given the established critical requirement for coupling D-galactan I O-PS export with polymerization? One possibility is the recognition of a conserved feature or property of the export substrate, and the only conserved element in the D-galactan I and polymannose O-PS export substrates is Und-PP–GlcNAc. However, the CPS model represents a rational alternative, where the coupling is potentially achieved by molecular recognition of proteins in the biosynthesis and export pathways rather than by any specific recognition of the glycan substrate. Nevertheless, the existence of different processes that couple polymerization and export of the various glycans does not necessarily mean that the actual mechanisms driving transport through the Wzm channel are different. Defining those mechanisms represents the next challenge.

Other O-PS NBDs from diverse pathogens of plants and humans are found in group D, and the evolutionary and physiological distinctions between group C and D NBDs are unclear. Three group D NBD representatives from *Y. enterocolitica* O-PS systems were included in the analysis. The O-PS structures from these bacteria are quite different in terms of structure. The O:3 antigen and O:9 antigens are homopolymers of 6-deoxy-L-atrose (47) and 4,6-dideoxy-4-formamido-D-mannopyranose (*N*-formyl perosamine) (19), respectively. In contrast, the O:5 O-PS contains a trisaccharide L-rhamnose backbone with substitutions of D-*threo*-

pent-2-ulose (xylulose) (48). The O:3 serotype deserves particular recognition as the first O-PS system for which an ABC transporter was identified (182). The *Yersinia* NBDs are relatively well separated within group D. In contrast, five NBDs from the genetic loci for the O-PSs of *Aggregatibacter* (*Actinobacillus*) *actinomycetemcomitans* and *Actinobacillus pleuropneumoniae* form a close-linked phylogenetic group, despite the structures of the corresponding O-PSs being remarkably different (4, 114, 115). The gene clusters for *A. actinomycetemcomitans* serotypes b, c, and f are highly conserved at the proximal and distal ends of the cluster, with a central region of lower G+C content that includes two to four genes which are unique to each serotype (64). This suggests that the serotypes evolved from a common ancestor and acquired the unique open reading frames from another source (99, 177). The gene cluster of *A. actinomycetemcomitans* serotype a is located in a different region of the chromosome from that for the other serotypes and contains several unique open reading frames, but relationships between the various ABC transporter components (40 to 60% amino acid identity) suggest that all of these systems share a common ancestor (155). It would be interesting to explore the exchangeability of the transporters between these serotypes. One would predict that like the case for *Klebsiella*, the NBDs would show no specificity for a particular repeat unit structure. In contrast to the NBD homologs in *A. actinomycetemcomitans* serotypes a, b, c, and f, the NBD homolog from serotype e contains an extended C-terminal domain and clusters with similar proteins in group A. Thus, the extent of phylogenetic linkage of the NBDs is not necessarily determined by structural diversity in the O-PS structures within the genus *Aggregatibacter*/*Actinobacillus*.

The identification of the *V. cholerae* O1 NBD in group D provides critical insight that helps to explain contradictory information concerning the biosynthesis and export of the O1 O-PS. Although only one glycan backbone is found in O1 serotypes, two subgroups (Inaba and Ogawa) are distinguished on the basis of O-PS serology. The glycan is a homopolymer consisting of α1,2-linked 4-amino-4,6-dideoxy-D-mannose (perosamine), and the serological differences result from the presence (Ogawa) or absence (Inaba) of terminal 2-*O*-methyl residues (55). Ogawa serotypes express both the methylated and nonmethylated versions of the O-PS, indicating that the chain termination residue added by RfbT (WbeT) (151) is not an absolute requirement for export via the ABC transporter. This situation resembles the multiply methylated O-PS from *R. etli* CE3 (described above) and differs from the *E. coli* O8/O9a paradigm, where chain termination and export are obligatorily coupled. The involvement of an NBD protein lacking a predicted substrate-binding domain in *V. cholerae* O1 predicts an export system that does not recognize nonreducing terminal modifications on the glycan, akin to the *K. pneumoniae* D-galactan I prototype, which has no apparent specificity for either the O-PS repeat unit or any terminating residues.

## A COMMON MECHANISM FOR THE EXPORT OF SOME GRAM-POSITIVE LIPOGLYCANS?

Group D also contains NBD homologs from Gram-positive bacteria and resembles group A in the sense that it includes representatives from different glycoconjugate systems that may be linked via a unified glycan assembly process. Unfortunately, the precise glycan substrates for the Gram-positive group D NBDs are not always clear. The NBD homolog (RfbE; also known as Rv3781) from *Mycobacterium tuberculosis* and its cognate TMD (RfbD; also known as Rv3783) are separated on the chromosome by the gene encoding GlfT1. GlfT1 is a bifunctional galactosyltransferase that adds the first two galactofuranose residues to an acceptor in the formation of the galactan core of arabinogalactan (12). GlfT2 subsequently adds up to 20 galactofuranosyl residues (92) generating the polyisoprenoid lipid-linked intermediate (decaprenol-PP-GlcNAc-Rha-Gal$f_n$). Further elaboration of the arabinogalactan molecule involves the addition of arabinosyl residues by the Emb and Aft proteins (13, 158). An NBD homolog similar to Rv3781 is found in *Corynebacterium glutamicum*, which also synthesizes an arabinogalactan (34). Based on our current understanding of related systems, the most likely role for this ABC transporter appears to be the export of the lipid intermediate containing the galactan structure prior to its arabinosylation. This proposal invokes a spatial separation of galactan formation (at the cytoplasmic face of the membrane) and arabinosylation (at the periplasmic face), but there are supporting precedents from other glycoconjugate assembly systems. The donor for the arabinosyltransferases is Dec-P-arabinofuranose (reviewed in references 13 and 158). In Gram-negative bacteria, the use of Und-P-linked donors (rather than the prevalent Und-PP) generally signals a periplasmic modification process, with the donors first being exported by a dedicated transporter (not an ABC transporter). Examples include the addition of 4-aminoarabinose to LPS lipid A (120) and the phage-encoded modification of O-PSs by glucosylation in species such as *Shigella* (51, 76). ArnE and ArnF, two small proteins of 111 and 128 amino acids, respectively, have been implicated in flipping the Und-P-4-aminoarabinose precursor involved in LPS modification (176). ArnE and ArnF each contain four predicted transmembrane helices and share both topological and sequence similarity with members of the drug/metabolite transporter superfamily, which includes EmrE from *E. coli* (176). GtrA is predicted to function in flipping Und-P-linked precursors required for O-PS modifications in *Shigella* (51). Like ArnE and ArnF, GtrA is a small protein (120 amino acids) with four transmembrane helices (76). Although their sequence similarity is low, GtrA may share structural similarity with EmrE (76). Interestingly, *M. tuberculosis* encodes a GtrA homolog (Rv3789), within the arabinogalactan biosynthesis locus, that may serve as the flippase for Dec-P-arabinose precursors. The hypothetical function for the Rv3783/Rv3781 ABC transporter in the export of the *M. tuberculosis* galactan remains to be tested, but there is significant interest in this area of glycobiology, as arabinogalactan is an essential part of the mycobacterial cell wall and is an excellent target for therapeutics. Indeed, the Emb proteins (arabinofuranosyl transferases) represent the target for the antimycobacterial agent ethambutol (11, 50).

## THE TEICHOIC ACID EXPORTERS—GROUP E NBDs

Group E represents a relatively tight cluster of NBDs involved in the export of the polyol phosphate teichoic acids in Gram-positive bacteria. While these molecules are clearly not oligo- or polysaccharides *per se*, their mode of synthesis is quite similar to those for many other bacterial glycoconjugates. It is

also important that components such as ribitol, glycerol, and phosphate are found occasionally in the repeat unit structures of CPSs and O-PSs. For example, all three components are found among the CPS structures of *Streptococcus pneumoniae*, and the genetic basis for the synthesis and diversity of these polymers has been established (1, 91). The best-characterized teichoic acid system is arguably the polyglycerol phosphate polymer from *B. subtilis* 168, where the cognate NBD is TagH (83). Notably, the TagH homologs from *B. subtilis* 168 and W23 are very similar, sharing 92% identity and 96% similarity, although the W23 teichoic acid is a polyribitol phosphate (6, 7, 170). It remains to be determined whether these transporters, like some ABC transporters with NBDs found in groups C, D, and F, do not discriminate between the structures of the teichoic acid backbones. In *B. subtilis* 168, chain-length regulation of polymers is apparently determined by the polymerase TagF, and the membrane association of TagF is important for this activity (135). The potential involvement of the teichoic acid ABC exporter in chain-length regulation has not been examined.

Group E contains NBD proteins of various lengths. Although some of them contain an extended C-terminal domain, topology predictions indicate that this domain is actually periplasmic and therefore unlikely to be involved in the recognition of teichoic acids for export from the cytoplasm (44). For example, the C-terminal regions of the TagH homologs from *Enterococcus faecalis* and *Lactococcus lactis* contain a predicted LysM domain following the transmembrane domain, which could be involved in binding to peptidoglycan, to which wall teichoic acids are covalently linked in the periplasm (117). The TagH homolog of *Bacillus anthracis* contains a predicted SH3b domain following a predicted transmembrane segment. SH3 domains are found in eukaryotes and viruses and are involved in protein-protein interactions (110, 172). SH3b domains are a subset of SH3 proteins found in prokaryotes and are involved in cell wall binding (89, 117, 172). The SH3b domain within the *B. anthracis* TagH homolog may also be involved in peptidoglycan binding in the periplasm. How these domains contribute to the overall assembly of teichoic acid is uncertain, but there is some evidence to suggest that the periplasmic C-terminal domain is not essential for viability of *B. subtilis* in the laboratory (83).

## HALF-EXPORTERS—NBDs IN GROUP G

Group G contains homologs of PglK and MsbA. These proteins lack additional C-terminal domains. The MsbA (LPS lipid A export) homologs generally form a relatively tight cluster, as one might expect given the highly conserved nature of the lipid A export substrate (120). The exception is the MsbA2 protein from *Bordetella petrii*, which clusters with the PglK (N-linked glycan export) homologs. One wonders if this representative is improperly annotated and plays a cellular role that is different from lipid A export.

The lipid A-core oligosaccharide portion of LPS is exported across the inner membrane by a process requiring the ABC transporter MsbA (120). While MsbA is widely assumed to be the lipid A exporter, and this interpretation is probably correct, it is important to remember that biochemical data verifying the ATP-dependent flipping of lipid A by MsbA are still lacking. The requirement for the first two Kdo residues of the inner core for cell viability in *E. coli* is a result of the inefficient export of lipid A without these residues (72). Addition of the first two residues of the core oligosaccharide is required prior to completion of lipid A acylation in *E. coli* K-12 and for efficient export of lipid A. Overexpression of MsbA is able to suppress the requirement for completion of the lipid A-$Kdo_2$ molecule for efficient export (65, 95). The possible involvement of MsbA in phospholipid export is less certain. Initial reports of MsbA depletion in *E. coli* suggested that MsbA might also be involved in phospholipid export (38, 39). However, other studies of an *msbA* mutant of *N. meningitidis* (an organism that does not require LPS for viability [145]) indicated that MsbA is not involved in phospholipid transport (160). It has been suggested that the accumulation of phospholipids observed in MsbA-depleted cells of *E. coli* may have been a secondary effect due to loss of LPS transport (160). Alternatively, the range of substrates for different homologs of MsbA may vary in different bacteria. It is certainly true that MsbA can export more than one substrate, as efflux of amphipathic drugs is recognized and the binding site(s) for these compounds is distinct from the high-affinity lipid A binding site (139). Furthermore, the multidrug transporter LmrA from *Lactobacillus lactis* can replace MsbA in mutant complementation (124). The recent discovery of the role of the multicomponent Mla system in retrograde transport of phospholipids to maintain the asymmetry in the outer membrane (90) represents a significant step forward in understanding lipid trafficking in Gram-negative bacteria. However, the picture is still incomplete, as the precise mechanism of phospholipid export remains to be identified.

PglK is encoded by the *pgl* (*p*rotein *gl*ycosylation) locus in *C. jejuni* NCTC 11168, and similar clusters have been found in other mucosal pathogens (157). The glycan product of the Pgl proteins is a branched, seven-residue glycan (178). Expression of the *C. jejuni pgl* locus in *E. coli* has facilitated study of the functions of the encoded gene products (166). Synthesis of Und-PP-linked glycan intermediates takes place at the cytoplasmic face of the inner membrane. The glycan is synthesized by sequential transfer of glycosyl residues by the glycosyltransferases (PglAHIJ) prior to its export across the inner membrane by the ABC transporter, PglK (2). Once exported, the glycan is ligated to specific consensus sites in the protein carriers by the oligosaccharyltransferase PglB (78). PglB exhibits a strict requirement for an acetamidosugar at the reducing end of the glycan (165) but has relaxed substrate specificity for the remaining sugars in the glycan (41, 150, 165). PglK is able to export O-PS repeat units from the *E. coli* O7 and O16 O-PSs in a chimeric experimental system. In the native *E. coli* system, the O7 and O16 Und-PP-linked intermediates are exported by Wzx, as part of a Wzx/Wzy-dependent O-PS pathway (2).

PglK seems to be an effective oligosaccharide transporter, suggesting that it behaves like Wzx in its transport abilities (2). It is not yet clear whether PglK can also export longer-chain Und-PP-linked polysaccharides, but interesting new insight into a PglK homolog (Wzk [also a group G protein]) suggests that this may be the case (58). Wzk is required for the export of O-PSs bearing the Lewis X and Lewis Y antigens, but it shows relaxed specificity with respect to the structure of the Und-PP-linked export substrates. Unlike *C. jejuni*, *H. pylori* does not contain the N-glyco-

sylation machinery. The identification of Wzk as the first half-transporter involved in O-PS export further highlights the close functional and evolutionary connections between the export systems used for different glycoconjugates.

## PHYLOGENETIC OUTLIERS

Based on the current analysis, three NBD proteins show no clear phylogenetic linkage to any of the seven groups. These may form part of a branch within the tree once more sequence data for more gene clusters become available. WbmM and WbmN from *Bordetella parapertussis* are encoded in the same O-PS biosynthesis locus (118). The O-PS of *B. parapertussis* terminates in a derivative of 2,3,4-triamino-2,3,4-trideoxy-L-galacturonamide (119). WbmM and WbmN both contain an extended C-terminal region, but the sequences differ and the roles of the two NBD homologs are not known. RfbB from the O-PS assembly system in *Myxococcus xanthus* (53) also contains an extended C-terminal domain. Just downstream of RfbB is a protein showing weak similarity to PMT (dolichyl phosphate-mannose-protein mannosyltransferase) proteins, raising the possibility that cell surface proteins in *M. xanthus* may be glycosylated with O-PS units. This situation has been well documented for some *P. aeruginosa* isolates, where repeat units of O-PS are used to O-glycosylate a pilin protein (20, 30). Until recently, this phenomenon was confined to O-PSs synthesized via a Wzx/Wzy-dependent process. However, the ABC transporter-dependent O-PS biosynthesis pathway (group D NBD) from *A. actinomycetemcomitans* serotype b can apparently be exploited in a similar fashion (159). Further work will be required to determine whether the predicted *M. xanthus* oligosaccharyltransferase is involved in protein glycosylation, but it could provide another example highlighting the flexibility of the biosynthesis machinery in generating polysaccharide substituents for a variety of ligands (i.e., lipids and proteins).

## CONCLUSIONS

The bioinformatic survey presented here provides intriguing insight into the structure-function relationships shared by NBDs from ABC transporters involved in the export of oligo- and polysaccharides. In all, seven phylogenetic groups of NBDs were identified, and to the extent that information on structure and biosynthesis is available, the NBD phylogeny is consistent with the biosynthetic mechanism. The majority of the ABC transporters involved in bacterial glycoconjugate assembly follow a format of independent TMD and NBD polypeptides. The exceptions are the half-transporters, such as MsbA and PglK, and it is likely that more examples will be found as the extent of available sequence data increases. Perhaps not surprisingly, these NBDs are well separated from the other representatives. Group F NBDs are from CPS assembly systems. The CPS repeat unit structures are diverse, but many are linked by the presence of diacylglycerol at the reducing terminus, and this may be the part of the molecule recognized by the exporter. If so, this would explain both their exchangeability and the close phylogenetic connections of ABC transporters within this group. The teichoic acid NBDs in group E also form a discrete group, which is perhaps not too surprising given the relatively limited range of teichoic acid structures. NBDs in groups C and D are implicated in export of a wide range of glycans in an array of bacteria. While the majority are Gram-negative O-PSs, the presence of an *M. tuberculosis* ABC transporter in this group raises interesting questions about arabinogalactan assembly. From the perspective of fundamental microbial physiology and biochemistry, the NBDs in groups A and B are particularly interesting. These groups include NBDs from diverse eubacteria (Gram-negative and Gram-positive) and archaea, reflecting a rich array of biology. All of the group A and B NBDs have extended C-terminal domains. Some are predicted to contain CBMs capable of binding the export substrate (potentially via nonreducing terminal residues), similar to prototype ABC-dependent O-PS systems. However, the observation that these NBDs have significantly more sequence diversity, with a wide variety of putative enzymatic activities contained within the C-terminal domains, was entirely unanticipated. These provide fascinating candidates for further study and may identify new concepts in glycoconjugate assembly. As new genomic information becomes available, analyses such as these can bring valuable initial insight into assembly processes and provide a template for further (directed) experiments.

## REFERENCES


1. **Aanensen, D. M., A. Mavroidi, S. D. Bentley, P. R. Reeves, and B. G. Spratt.** 2007. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. J. Bacteriol. **189:**7856–7876.
2. **Alaimo, C., I. Catrein, L. Morf, C. L. Marolda, N. Callewaert, M. A. Valvano, M. F. Feldman, and M. Aebi.** 2006. Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. EMBO J. **25:**967–976.
3. **Allison, G. E., and N. K. Verma.** 2000. Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. Trends Microbiol. **8:**17–23.
4. **Altman, E., J. R. Brisson, and M. B. Perry.** 1986. Structure of the O-chain of the lipopolysaccharide of *Haemophilus pleuropneumoniae* serotype 1. Biochem. Cell Biol. **64:**1317–1325.
5. **Andreishcheva, E. N., and W. F. Vann.** 2006. Gene products required for *de novo* synthesis of polysialic acid in *Escherichia coli* K1. J. Bacteriol. **188:**1786–1797.
6. **Araki, Y., and E. Ito.** 1989. Linkage units in cell walls of gram-positive bacteria. Crit. Rev. Microbiol. **17:**121–135.
7. **Armstrong, J. J., J. Baddiley, and J. G. Buchanan.** 1960. Structure of the ribitol teichoic acid from the walls of *Bacillus subtilis*. Biochem. J. **76:**610–621.
8. **Arrecubieta, C., T. C. Hammarton, B. Barrett, S. Chareonsudjai, N. Hodson, D. Rainey, and I. S. Roberts.** 2001. The transport of group 2 capsular polysaccharides across the periplasmic space in *Escherichia coli*. Roles for the KpsE and KpsD proteins. J. Biol. Chem. **276:**4245–4250.
9. **Arsenault, T. L., D. W. Hughes, D. B. MacLean, W. A. Szarek, A. M. B. Kropinski, and J. S. Lam.** 1991. Structural studies on the polysaccharide portion of "A-band" lipopolysaccharide from a mutant (AK1401) of *Pseudomonas aeruginosa* strain PAO1. Can. J. Chem. **69:**1273–1280.
10. **Bacon, D. J., C. M. Szymanski, D. H. Burr, R. P. Silver, R. A. Alm, and P. Guerry.** 2001. A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. Mol. Microbiol. **40:**769–777.
11. **Belanger, A. E., G. S. Besra, M. E. Ford, K. Mikusova, J. T. Belisle, P. J. Brennan, and J. M. Inamine.** 1996. The *embAB* genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. Proc. Natl. Acad. Sci. U. S. A. **93:**11919–11924.
12. **Belanova, M., P. Dianiskova, P. J. Brennan, G. C. Completo, N. L. Rose, T. L. Lowary, and K. Mikusova.** 2008. Galactosyl transferases in mycobacterial cell wall synthesis. J. Bacteriol. **190:**1141–1145.
13. **Berg, S., D. Kaur, M. Jackson, and P. J. Brennan.** 2007. The glycosyltransferases of *Mycobacterium tuberculosis*—roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. Glycobiology **17:**35R–56R.
14. **Biemans-Oldehinkel, E., M. K. Doeven, and B. Poolman.** 2006. ABC transporter architecture and regulatory roles of accessory domains. FEBS Lett. **580:**1023–1035.
15. **Bliss, J. M., and R. P. Silver.** 1996. Coating the surface: a model for expression of capsular polysialic acid in *Escherichia coli* K1. Mol. Microbiol. **21:**221–231.
16. **Bourdineaud, J. P., J. J. Bono, R. Ranjeva, and J. V. Cullimore.** 1995.

Enzymatic radiolabelling to a high specific activity of legume lipo-oligosaccharidic nodulation factors from *Rhizobium meliloti*. Biochem. J. **306:**259–264.
17. **Bronner, D., V. Sieberth, C. Pazzani, A. Smith, G. Boulnois, I. Roberts, B. Jann, and K. Jann.** 1993. Synthesis of the K5 (group II) capsular polysaccharide in transport-deficient recombinant *Escherichia coli*. FEMS Microbiol. Lett. **113:**279–284.
18. **Burger, M. M., and L. Glaser.** 1964. The synthesis of teichoic acids. I. Polyglycerophosphate. J. Biol. Chem. **239:**3168–3177.
19. **Caroff, M., D. R. Bundle, and M. B. Perry.** 1984. Structure of the O-chain of the phenol-phase soluble cellular lipopolysaccharide of *Yersinia enterocolitica* serotype O:9. Eur. J. Biochem. **139:**195–200.
20. **Castric, P., F. J. Cassels, and R. W. Carlson.** 2001. Structural characterization of the *Pseudomonas aeruginosa* 1244 pilin glycan. J. Biol. Chem. **276:**26479–26485.
21. **Chaban, B., S. Voisin, J. Kelly, S. M. Logan, and K. F. Jarrell.** 2006. Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae N*-linked glycans: insight into *N*-linked glycosylation pathways in archaea. Mol. Microbiol. **61:**259–268.
22. **Chart, H., D. H. Shaw, E. E. Ishiguro, and T. J. Trust.** 1984. Structural and immunochemical homogeneity of *Aeromonas salmonicida* lipopolysaccharide. J. Bacteriol. **158:**16–22.
23. **Chen, J., G. Lu, J. Lin, A. L. Davidson, and F. A. Quiocho.** 2003. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. Mol. Cell **12:**651–661.
24. **Chu, S., B. Noonan, S. Cavaignac, and T. J. Trust.** 1995. Endogenous mutagenesis by an insertion sequence element identifies *Aeromonas salmonicida* AbcA as an ATP-binding cassette transport protein required for biogenesis of smooth lipopolysaccharide. Proc. Natl. Acad. Sci. U. S. A. **92:**5754–5758.
25. **Chu, S., and T. J. Trust.** 1993. An *Aeromonas salmonicida* gene which influences A-protein expression in *Escherichia coli* encodes a protein containing an ATP-binding cassette and maps beside the surface array protein gene. J. Bacteriol. **175:**3105–3114.
26. **Clarke, B. R., L. Cuthbertson, and C. Whitfield.** 2004. Nonreducing terminal modifications determine the chain length of polymannose O antigens of *Escherichia coli* and couple chain termination to polymer export via an ATP-binding cassette transporter. J. Biol. Chem. **279:**35709–35718.
27. **Clarke, B. R., L. K. Greenfield, C. Bouwman, and C. Whitfield.** 2009. Coordination of polymerization, chain termination, and export in assembly of the *Escherichia coli* lipopolysaccharide O9a antigen in an ATP-binding cassette transporter-dependent pathway. J. Biol. Chem. **284:**30662–30672.
28. **Clarke, B. R., and C. Whitfield.** 1992. Molecular cloning of the *rfb* region of *Klebsiella pneumoniae* serotype O1:K20: the *rfb* gene cluster is responsible for synthesis of the D-galactan I O polysaccharide. J. Bacteriol. **174:**4614–4621.
29. **Collins, R. F., K. Beis, C. Dong, C. H. Botting, C. McDonnell, R. C. Ford, B. R. Clarke, C. Whitfield, and J. H. Naismith.** 2007. The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **104:**2390–2395.
30. **Comer, J. E., M. A. Marshall, V. J. Blanch, C. D. Deal, and P. Castric.** 2002. Identification of the *Pseudomonas aeruginosa* 1244 pilin glycosylation site. Infect. Immun. **70:**2837–2845.
31. **Cuthbertson, L., M. S. Kimber, and C. Whitfield.** 2007. Substrate binding by a bacterial ABC transporter involved in polysaccharide export. Proc. Natl. Acad. Sci. U. S. A. **104:**19529–19534.
32. **Cuthbertson, L., I. L. Mainprize, J. H. Naismith, and C. Whitfield.** 2009. Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. Microbiol. Mol. Biol. Rev. **73:**155–177.
33. **Cuthbertson, L., J. Powers, and C. Whitfield.** 2005. The C-terminal domain of the nucleotide-binding domain protein Wzt determines substrate specificity in the ATP-binding cassette transporter for the lipopolysaccharide O-antigens in *Escherichia coli* serotypes O8 and O9a. J. Biol. Chem. **280:**30310–30319.
34. **Daffe, M., M. McNeil, and P. J. Brennan.** 1993. Major structural features of the cell wall arabinogalactans of *Mycobacterium*, *Rhodococcus*, and *Nocardia* spp. Carbohydr. Res. **249:**383–398.
35. **Dawson, R. J., and K. P. Locher.** 2006. Structure of a bacterial multidrug ABC transporter. Nature **443:**180–185.
36. **Dawson, R. J., and K. P. Locher.** 2007. Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. FEBS Lett. **581:**935–938.
37. **Diederichs, K., J. Diez, G. Greller, C. Muller, J. Breed, C. Schnell, C. Vonrhein, W. Boos, and W. Welte.** 2000. Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*. EMBO J. **19:**5951–5961.
38. **Doerrler, W. T., H. S. Gibbons, and C. R. Raetz.** 2004. MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. J. Biol. Chem. **279:**45102–45109.
39. **Doerrler, W. T., M. C. Reedy, and C. R. Raetz.** 2001. An *Escherichia coli* mutant defective in lipid export. J. Biol. Chem. **276:**11461–11464.
40. **Ehrhardt, D. W., E. M. Atkinson, K. F. Faull, D. I. Freedberg, D. P. Sutherlin, R. Armstrong, and S. R. Long.** 1995. *In vitro* sulfotransferase activity of NodH, a nodulation protein of *Rhizobium meliloti* required for host-specific nodulation. J. Bacteriol. **177:**6237–6245.
41. **Feldman, M. F., M. Wacker, M. Hernandez, P. G. Hitchen, C. L. Marolda, M. Kowarik, H. R. Morris, A. Dell, M. A. Valvano, and M. Aebi.** 2005. Engineering *N*-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **102:**3016–3021.
42. **Felsenstein, J.** 1989. Mathematics vs. evolution: mathematical evolutionary theory. Science **246:**941–942.
43. **Finke, A., D. Bronner, A. V. Nikolaev, B. Jann, and K. Jann.** 1991. Biosynthesis of the *Escherichia coli* K5 polysaccharide, a representative of group II capsular polysaccharides: polymerization *in vitro* and characterization of the product. J. Bacteriol. **173:**4088–4094.
44. **Formstone, A., R. Carballido-Lopez, P. Noirot, J. Errington, and D. J. Scheffers.** 2008. Localization and interactions of teichoic acid synthetic enzymes in *Bacillus subtilis*. J. Bacteriol. **190:**1812–1821.
45. **Forsberg, L. S., U. R. Bhat, and R. W. Carlson.** 2000. Structural characterization of the O-antigenic polysaccharide of the lipopolysaccharide from *Rhizobium etli* strain CE3. A unique O-acetylated glycan of discrete size, containing 3-O-methyl-6-deoxy-L-talose and 2,3,4-tri-O-, methyl-L fucose. J. Biol. Chem. **275:**18851–18863.
46. **Goetz, B. A., E. Perozo, and K. P. Locher.** 2009. Distinct gate conformations of the ABC transporter BtuCD revealed by electron spin resonance spectroscopy and chemical cross-linking. FEBS Lett. **583:**266–270.
47. **Gorshkova, R. P., E. N. Kalmykova, V. V. Isakov, and Y. S. Ovodov.** 1985. Structural studies on O-specific polysaccharides of lipopolysaccharides from *Yersinia enterocolitica* serovars O:1,2a,3, O:2a,2b,3 and O:3. Eur. J. Biochem. **150:**527–531.
48. **Gorshkova, R. P., E. N. Kalmykova, V. V. Isakov, and Y. S. Ovodov.** 1986. Structural studies on O-specific polysaccharides of lipopolysaccharides from *Yersinia enterocolitica* serovars O:5 and O:5,27. Eur. J. Biochem. **156:**391–397.
49. **Gotschlich, E. C., B. A. Fraser, O. Nishimura, J. B. Robbins, and T. Y. Liu.** 1981. Lipid on capsular polysaccharides of gram-negative bacteria. J. Biol. Chem. **256:**8915–8921.
50. **Goude, R., A. G. Amin, D. Chatterjee, and T. Parish.** 2009. The EmbC arabinosyltransferase is inhibited by ethambutol in *Mycobacterium tuberculosis*. Antimicrob. Agents Chemother. **53:**4138–4146.
51. **Guan, S., D. A. Bastin, and N. K. Verma.** 1999. Functional analysis of the O antigen glucosylation gene cluster of *Shigella flexneri* bacteriophage SfX. Microbiology **145:**1263–1273.
52. **Guan, S., A. J. Clarke, and C. Whitfield.** 2001. Functional analysis of the galactosyltransferases required for biosynthesis of D-galactan I, a component of the lipopolysaccharide O1 antigen of *Klebsiella pneumoniae*. J. Bacteriol. **183:**3318–3327.
53. **Guo, D., M. G. Bowden, R. Pershad, and H. B. Kaplan.** 1996. The *Myxococcus xanthus rfbABC* operon encodes an ATP-binding cassette transporter homolog required for O-antigen biosynthesis and multicellular development. J. Bacteriol. **178:**1631–1639.
54. **Hashimoto, H.** 2006. Recent structural studies of carbohydrate-binding modules. Cell. Mol. Life Sci. **63:**2954–2967.
55. **Hisatsune, K., S. Kondo, T. Iguchi, T. Ito, and K. Hiramatsu.** 1996. Lipopolysaccharides of *Escherichia coli* K12 strains that express cloned genes for the Ogawa and Inaba antigens of *Vibrio cholerae* O1: identification of O-antigenic factors. Microbiol. Immunol. **40:**621–626.
56. **Hoffman, J., B. Lindberg, and R. R. Brubaker.** 1980. Structural studies of the O-specific side-chains of the lipopolysaccharide from *Yersinia enterocolitica* Ye 128. Carbohydr. Res. **78:**212–214.
57. **Hollenstein, K., R. J. Dawson, and K. P. Locher.** 2007. Structure and mechanism of ABC transporter proteins. Curr. Opin. Struct. Biol. **17:**412–418.
58. **Hug, I., M. R. Couturier, M. M. Rooker, D. E. Taylor, M. Stein, and M. F. Feldman.** 2010. *Helicobacter pylori* lipopolysaccharide is synthesized via a novel pathway with an evolutionary connection to protein *N*-glycosylation. PLoS Pathog. **6:**e1000819. doi:10.1371/journal.ppat.1000819.
59. **Huson, D. H., and D. Bryant.** 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23:**254–267.
60. **Itoh, Y., J. D. Rice, C. Goller, A. Pannuri, J. Taylor, J. Meisner, T. J. Beveridge, J. F. Preston III, and T. Romeo.** 2008. Roles of *pgaABCD* genes in synthesis, modification, and export of the *Escherichia coli* biofilm adhesin poly-beta-1,6-*N*-acetyl-D-glucosamine. J. Bacteriol. **190:**3670–3680.
61. **Jann, K., G. Goldemann, C. Weisgerber, C. Wolf-Ullisch, and S. Kanegasaki.** 1982. Biosynthesis of the O9 antigen of *Escherichia coli*. Initial reaction and overall mechanism. Eur. J. Biochem. **127:**157–164.
62. **Jones, P. M., M. L. O'Mara, and A. M. George.** 2009. ABC transporters: a riddle wrapped in a mystery inside an enigma. Trends Biochem. Sci. **34:**520–531.
63. **Kahlig, H., D. Kolarich, S. Zayni, A. Scheberl, P. Kosma, C. Schaffer, and**

**P. Messner.** 2005. N-acetylmuramic acid as capping element of alpha-D-fucose-containing S-layer glycoprotein glycans from *Geobacillus tepidamans* GS5-97T. J. Biol. Chem. **280:**20292–20299.
64. **Kaplan, J. B., M. B. Perry, L. L. MacLean, D. Furgang, M. E. Wilson, and D. H. Fine.** 2001. Structural and genetic analyses of O polysaccharide from *Actinobacillus actinomycetemcomitans* serotype f. Infect. Immun. **69:**5375–5384.
65. **Karow, M., and C. Georgopoulos.** 1993. The essential *Escherichia coli msbA* gene, a multicopy suppressor of null mutations in the *htrB* gene, is related to the universally conserved family of ATP-dependent translocators. Mol. Microbiol. **7:**69–79.
66. **Keenleyside, W. J., and C. Whitfield.** 1996. A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* serovar *Borreze*. J. Biol. Chem. **271:**28581–28592.
67. **Kelly, R. F., W. B. Severn, J. C. Richards, M. B. Perry, L. L. MacLean, J. M. Tomas, S. Merino, and C. Whitfield.** 1993. Structural variation in the O-specific polysaccharides of *Klebsiella pneumoniae* serotype O1 and O8 lipopolysaccharide: evidence for clonal diversity in *rfb* genes. Mol. Microbiol. **10:**615–625.
68. **Kerr, I. D.** 2002. Structure and association of ATP-binding cassette transporter nucleotide-binding domains. Biochim. Biophys. Acta **1561:**47–64.
69. **Khare, D., M. L. Oldham, C. Orelle, A. L. Davidson, and J. Chen.** 2009. Alternating access in maltose transporter mediated by rigid-body rotations. Mol. Cell **33:**528–536.
70. **Kido, N., V. I. Torgov, T. Sugiyama, K. Uchiya, H. Sugihara, T. Komatsu, N. Kato, and K. Jann.** 1995. Expression of the O9 polysaccharide of *Escherichia coli*: sequencing of the *E. coli* O9 *rfb* gene cluster, characterization of mannosyl transferases, and evidence for an ATP-binding cassette transport system. J. Bacteriol. **177:**2178–2187.
71. **King, J. D., K. K. Poon, N. A. Webb, E. M. Anderson, D. J. McNally, J. R. Brisson, P. Messner, R. M. Garavito, and J. S. Lam.** 2009. The structural basis for catalytic function of GMD and RMD, two closely related enzymes from the GDP-D-rhamnose biosynthesis pathway. FEBS J. **276:**2686–2700.
72. **Klein, G., B. Lindner, W. Brabetz, H. Brade, and S. Raina.** 2009. *Escherichia coli* K-12 suppressor-free mutants lacking early glycosyltransferases and late acyltransferases: minimal lipopolysaccharide structure and induction of envelope stress response. J. Biol. Chem. **284:**15369–15389.
73. **Kol, O., J. M. Wieruszeski, G. Strecker, B. Fournet, R. Zalisz, and P. Smets.** 1992. Structure of the O-specific polysaccharide chain of *Klebsiella pneumoniae* O1K2 (NCTC 5055) lipopolysaccharide. A complementary elucidation. Carbohydr. Res. **236:**339–344.
74. **Kol, O., J. M. Wieruszeski, G. Strecker, J. Montreuil, B. Fournet, R. Zalisz, and P. Smets.** 1991. Structure of the O-specific polysaccharide chain from *Klebsiella pneumoniae* O1K2 (NCTC 5055) lipopolysaccharide. Carbohydr. Res. **217:**117–125.
75. **Koplin, R., J. R. Brisson, and C. Whitfield.** 1997. UDP-galactofuranose precursor required for formation of the lipopolysaccharide O antigen of *Klebsiella pneumoniae* serotype O1 is synthesized by the product of the *rfbD*KPO1 gene. J. Biol. Chem. **272:**4121–4128.
76. **Korres, H., M. Mavris, R. Morona, P. A. Manning, and N. K. Verma.** 2005. Topological analysis of GtrA and GtrB proteins encoded by the serotype-converting cassette of *Shigella flexneri*. Biochem. Biophys. Res. Commun. **328:**1252–1260.
77. **Kos, V., L. Cuthbertson, and C. Whitfield.** 2009. The *Klebsiella pneumoniae* O2a antigen defines a second mechanism for O antigen ATP-binding cassette transporters. J. Biol. Chem. **284:**2947–2956.
78. **Kowarik, M., N. M. Young, S. Numao, B. L. Schulz, I. Hug, N. Callewaert, D. C. Mills, D. C. Watson, M. Hernandez, J. F. Kelly, M. Wacker, and M. Aebi.** 2006. Definition of the bacterial *N*-glycosylation site consensus sequence. EMBO J. **25:**1957–1966.
79. **Kroll, J. S., B. Loynds, L. N. Brophy, and E. R. Moxon.** 1990. The *bex* locus in encapsulated *Haemophilus influenzae*: a chromosomal region involved in capsule polysaccharide export. Mol. Microbiol. **4:**1853–1862.
80. **Kroncke, K. D., G. Boulnois, I. Roberts, D. Bitter-Suermann, J. R. Golecki, B. Jann, and K. Jann.** 1990. Expression of the *Escherichia coli* K5 capsular antigen: immunoelectron microscopic and biochemical studies with recombinant *E. coli*. J. Bacteriol. **172:**1085–1091.
81. **Kroncke, K. D., I. Orskov, F. Orskov, B. Jann, and K. Jann.** 1990. Electron microscopic study of coexpression of adhesive protein capsules and polysaccharide capsules in *Escherichia coli*. Infect. Immun. **58:**2710–2714.
82. **Laeremans, T., I. Caluwaerts, C. Verreth, M. A. Rogel, J. Vanderleyden, and E. Martinez-Romero.** 1996. Isolation and characterization of *Rhizobium tropici* Nod factor sulfation genes. Mol. Plant Microbe Interact. **9:**492–500.
83. **Lazarevic, V., and D. Karamata.** 1995. The *tagGH* operon of *Bacillus subtilis* 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids. Mol. Microbiol. **16:**345–355.
84. **Lerouge, I., T. Laeremans, C. Verreth, J. Vanderleyden, C. Van Soom, A. Tobin, and R. W. Carlson.** 2001. Identification of an ATP-binding cassette transporter for export of the O-antigen across the inner membrane in *Rhizobium etli* based on the genetic, functional, and structural analysis of an LPS mutant deficient in O-antigen. J. Biol. Chem. **276:**17190–17198.
85. **Lerouge, I., C. Verreth, J. Michiels, R. W. Carlson, A. Datta, M. Y. Gao, and J. Vanderleyden.** 2003. Three genes encoding for putative methyl- and acetyltransferases map adjacent to the *wzm* and *wzt* genes and are essential for O-antigen biosynthesis in *Rhizobium etli* CE3. Mol. Plant Microbe Interact. **16:**1085–1093.
86. **Lindberg, B., J. Lonngren, and W. Nimmich.** 1972. Structural studies on *Klebsiella* O group 5 lipopolysaccharides. Acta Chem. Scand. **26:**2231–2236.
87. **Lo, R. Y., L. J. McKerral, T. L. Hills, and M. Kostrzynska.** 2001. Analysis of the capsule biosynthetic locus of *Mannheimia* (*Pasteurella*) *haemolytica* A1 and proposal of a nomenclature system. Infect. Immun. **69:**4458–4464.
88. **Locher, K. P., A. T. Lee, and D. C. Rees.** 2002. The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. Science **296:**1091–1098.
89. **Lu, J. Z., T. Fujiwara, H. Komatsuzawa, M. Sugai, and J. Sakon.** 2006. Cell wall-targeting domain of glycylglycine endopeptidase distinguishes among peptidoglycan cross-bridges. J. Biol. Chem. **281:**549–558.
90. **Malinverni, J. C., and T. J. Silhavy.** 2009. An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. Proc. Natl. Acad. Sci. U. S. A. **106:**8009–8014.
91. **Mavroidi, A., D. M. Aanensen, D. Godoy, I. C. Skovsted, M. S. Kaltoft, P. R. Reeves, S. D. Bentley, and B. G. Spratt.** 2007. Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. J. Bacteriol. **189:**7841–7855.
92. **May, J. F., R. A. Splain, C. Brotschi, and L. L. Kiessling.** 2009. A tethering mechanism for length control in a processive carbohydrate polymerization. Proc. Natl. Acad. Sci. U. S. A. **106:**11851–11856.
93. **McGuire, E. J., and S. B. Binkley.** 1964. The structure and chemistry of colominic acid. Biochemistry **3:**247–251.
94. **McNulty, C., J. Thompson, B. Barrett, L. Lord, C. Andersen, and I. S. Roberts.** 2006. The cell surface expression of group 2 capsular polysaccharides in *Escherichia coli*: the role of KpsD, RhsA and a multi-protein complex at the pole of the cell. Mol. Microbiol. **59:**907–922.
95. **Meredith, T. C., P. Aggarwal, U. Mamat, B. Lindner, and R. W. Woodard.** 2006. Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. ACS Chem. Biol. **1:**33–42.
96. **Morona, R., L. Purins, A. Tocilj, A. Matte, and M. Cygler.** 2009. Sequence-structure relationships in polysaccharide co-polymerase (PCP) proteins. Trends Biochem. Sci. **34:**78–84.
97. **Mourez, M., M. Hofnung, and E. Dassa.** 1997. Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits. EMBO J. **16:**3066–3077.
98. **Muller, M. G., L. S. Forsberg, and D. H. Keating.** 2009. The *rkp-1* cluster is required for secretion of Kdo homopolymeric capsular polysaccharide in *Sinorhizobium meliloti* strain Rm1021. J. Bacteriol. **191:**6988–7000.
99. **Nakano, Y., Y. Yoshida, N. Suzuki, Y. Yamashita, and T. Koga.** 2000. A gene cluster for the synthesis of serotype d-specific polysaccharide antigen in *Actinobacillus actinomycetemcomitans*. Biochim. Biophys. Acta **1493:**259–263.
100. **Noonan, B., and T. J. Trust.** 1995. The leucine zipper of *Aeromonas salmonicida* AbcA is required for the transcriptional activation of the P2 promoter of the surface-layer structural gene, *vapA*, in *Escherichia coli*. Mol. Microbiol. **17:**379–386.
101. **Noonan, B., and T. J. Trust.** 1997. The synthesis, secretion and role in virulence of the paracrystalline surface protein layers of *Aeromonas salmonicida* and *A. hydrophila*. FEMS Microbiol. Lett. **154:**1–7.
102. **Novotny, R., A. Pfoestl, P. Messner, and C. Schaffer.** 2004. Genetic organization of chromosomal S-layer glycan biosynthesis loci of Bacillaceae. Glycoconj. J. **20:**435–447.
103. **Oancea, G., M. L. O'Mara, W. F. Bennett, D. P. Tieleman, R. Abele, and R. Tampe.** 2009. Structural arrangement of the transmission interface in the antigen ABC transport complex TAP. Proc. Natl. Acad. Sci. U. S. A. **106:**5551–5556.
104. **Ojeda, K. J., J. M. Box, and K. D. Noel.** 2009. Genetic basis for *Rhizobium etli* CE3 O-antigen O-methylated residues that vary according to growth conditions. J. Bacteriol. **192:**679–690.
105. **Oldham, M. L., D. Khare, F. A. Quiocho, A. L. Davidson, and J. Chen.** 2007. Crystal structure of a catalytic intermediate of the maltose transporter. Nature **450:**515–521.
106. **Orelle, C., T. Ayvaz, R. M. Everly, C. S. Klug, and A. L. Davidson.** 2008. Both maltose-binding protein and ATP are required for nucleotide-binding domain closure in the intact maltose ABC transporter. Proc. Natl. Acad. Sci. U. S. A. **105:**12837–12842.
107. **Oxley, D., and S. G. Wilkinson.** 1989. Structures of neutral glycans isolated from the lipopolysaccharides of reference strains for *Serratia marcescens* serogroups O16 and O20. Carbohydr. Res. **193:**241–248.
108. **Pavelka, M. S., Jr., S. F. Hayes, and R. P. Silver.** 1994. Characterization of KpsT, the ATP-binding component of the ABC-transporter involved with the export of capsular polysialic acid in *Escherichia coli* K1. J. Biol. Chem. **269:**20149–20158.
109. **Pavelka, M. S., Jr., L. F. Wright, and R. P. Silver.** 1991. Identification of

two genes, *kpsM* and *kpsT*, in region 3 of the polysialic acid gene cluster of *Escherichia coli* K1. J. Bacteriol. **173:**4603–4610.
110. **Pawson, T.** 1994. SH2 and SH3 domains in signal transduction. Adv. Cancer Res. **64:**87–110.
111. **Pazzani, C., C. Rosenow, G. J. Boulnois, D. Bronner, K. Jann, and I. S. Roberts.** 1993. Molecular analysis of region 1 of the *Escherichia coli* K5 antigen gene cluster: a region encoding proteins involved in cell surface expression of capsular polysaccharide. J. Bacteriol. **175:**5978–5983.
112. **Pearce, R., and I. S. Roberts.** 1995. Cloning and analysis of gene clusters for production of the *Escherichia coli* K10 and K54 antigens: identification of a new group of *serA*-linked capsule gene clusters. J. Bacteriol. **177:**3992–3997.
113. **Perepelov, A. V., D. Li, B. Liu, S. N. Senchenkova, D. Guo, S. D. Shevelev, A. S. Shashkov, X. Guo, L. Feng, Y. A. Knirel, and L. Wang.** 2009. Structural and genetic characterization of *Escherichia coli* O99 antigen. FEMS Immunol. Med. Microbiol. **57:**80–87.
114. **Perry, M. B., L. L. MacLean, R. Gmur, and M. E. Wilson.** 1996. Characterization of the O-polysaccharide structure of lipopolysaccharide from *Actinobacillus actinomycetemcomitans* serotype b. Infect. Immun. **64:**1215–1219.
115. **Perry, M. B., L. M. MacLean, J. R. Brisson, and M. E. Wilson.** 1996. Structures of the antigenic O-polysaccharides of lipopolysaccharides produced by *Actinobacillus actinomycetemcomitans* serotypes a, c, d and e. Eur. J. Biochem. **242:**682–688.
116. **Pigeon, R. P., and R. P. Silver.** 1994. Topological and mutational analysis of KpsM, the hydrophobic component of the ABC-transporter involved in the export of polysialic acid in *Escherichia coli* K1. Mol. Microbiol. **14:**871–881.
117. **Ponting, C. P., L. Aravind, J. Schultz, P. Bork, and E. V. Koonin.** 1999. Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. J. Mol. Biol. **289:**729–745.
118. **Preston, A., A. G. Allen, J. Cadisch, R. Thomas, K. Stevens, C. M. Churcher, K. L. Badcock, J. Parkhill, B. Barrell, and D. J. Maskell.** 1999. Genetic basis for lipopolysaccharide O-antigen biosynthesis in bordetellae. Infect. Immun. **67:**3763–3767.
119. **Preston, A., B. O. Petersen, J. O. Duus, J. Kubler-Kielb, G. Ben-Menachem, J. Li, and E. Vinogradov.** 2006. Complete structures of *Bordetella bronchiseptica* and *Bordetella parapertussis* lipopolysaccharides. J. Biol. Chem. **281:**18135–18144.
120. **Raetz, C. R., C. M. Reynolds, M. S. Trent, and R. E. Bishop.** 2007. Lipid A modification systems in gram-negative bacteria. Annu. Rev. Biochem. **76:**295–329.
121. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu. Rev. Biochem. **71:**635–700.
122. **Rees, D. C., E. Johnson, and O. Lewinson.** 2009. ABC transporters: the power to change. Nat. Rev. Mol. Cell. Biol. **10:**218–227.
123. **Remminghorst, U., and B. H. Rehm.** 2006. Bacterial alginates: from biosynthesis to applications. Biotechnol. Lett. **28:**1701–1712.
124. **Reuter, G., T. Janvilisri, H. Venter, S. Shahi, L. Balakrishnan, and H. W. van Veen.** 2003. The ATP binding cassette multidrug transporter LmrA and lipid transporter MsbA have overlapping substrate specificities. J. Biol. Chem. **278:**35193–35198.
125. **Reyes, C. L., and G. Chang.** 2005. Structure of the ABC transporter MsbA in complex with ADP.vanadate and lipopolysaccharide. Science **308:**1028–1031.
126. **Rick, P. D., G. L. Hubbard, and K. Barr.** 1994. Role of the *rfe* gene in the synthesis of the O8 antigen in *Escherichia coli* K-12. J. Bacteriol. **176:**2877–2884.
127. **Rigg, G. P., B. Barrett, and I. S. Roberts.** 1998. The localization of KpsC, S and T, and KfiA, C and D proteins involved in the biosynthesis of the *Escherichia coli* K5 capsular polysaccharide: evidence for a membrane-bound complex. Microbiology **144:**2905–2914.
128. **Roberts, I., R. Mountford, N. High, D. Bitter-Suermann, K. Jann, K. Timmis, and G. Boulnois.** 1986. Molecular cloning and analysis of genes for production of K5, K7, K12, and K92 capsular polysaccharides in *Escherichia coli*. J. Bacteriol. **168:**1228–1233.
129. **Roberts, M., I. Roberts, T. K. Korhonen, K. Jann, D. Bitter-Suermann, G. J. Boulnois, and P. H. Williams.** 1988. DNA probes for K-antigen (capsule) typing of *Escherichia coli*. J. Clin. Microbiol. **26:**385–387.
130. **Romling, U.** 2002. Molecular biology of cellulose production in bacteria. Res. Microbiol. **153:**205–212.
131. **Schaffer, C., and P. Messner.** 2005. The structure of secondary cell wall polymers: how Gram-positive bacteria stick their cell walls together. Microbiology **151:**643–651.
132. **Schaffer, C., N. Muller, R. Christian, M. Graninger, T. Wugeditsch, A. Scheberl, and P. Messner.** 1999. Complete glycan structure of the S-layer glycoprotein of *Aneurinibacillus thermoaerophilus* GS4-97. Glycobiology **9:**407–414.
133. **Schaffer, C., T. Wugeditsch, H. Kahlig, A. Scheberl, S. Zayni, and P. Messner.** 2002. The surface layer (S-layer) glycoprotein of *Geobacillus stearothermophilus* NRS 2004/3a. Analysis of its glycosylation. J. Biol. Chem. **277:**6230–6239.
134. **Schaffer, C., T. Wugeditsch, C. Neuninger, and P. Messner.** 1996. Are S-layer glycoproteins and lipopolysaccharides related? Microb. Drug Resist. **2:**17–23.
135. **Schertzer, J. W., and E. D. Brown.** 2008. Use of CDP-glycerol as an alternate acceptor for the teichoic acid polymerase reveals that membrane association regulates polymer length. J. Bacteriol. **190:**6940–6947.
136. **Schmitt, L., H. Benabdelhak, M. A. Blight, I. B. Holland, and M. T. Stubbs.** 2003. Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. J. Mol. Biol. **330:**333–342.
137. **Schultze, M., C. Staehelin, H. Rohrig, M. John, J. Schmidt, E. Kondorosi, J. Schell, and A. Kondorosi.** 1995. *In vitro* sulfotransferase activity of *Rhizobium meliloti* NodH protein: lipochitooligosaccharide nodulation signals are sulfated after synthesis of the core structure. Proc. Natl. Acad. Sci. U. S. A. **92:**2706–2709.
138. **Senior, A. E., M. K. al-Shawi, and I. L. Urbatsch.** 1995. The catalytic cycle of P-glycoprotein. FEBS Lett. **377:**285–289.
139. **Siarheyeva, A., and F. J. Sharom.** 2009. The ABC transporter MsbA interacts with lipid A and amphipathic drugs at different sites. Biochem. J. **419:**317–328.
140. **Silver, R. P., W. Aaronson, and W. F. Vann.** 1987. Translocation of capsular polysaccharides in pathogenic strains of *Escherichia coli* requires a 60-kilodalton periplasmic protein. J. Bacteriol. **169:**5489–5495.
141. **Silver, R. P., C. W. Finn, W. F. Vann, W. Aaronson, R. Schneerson, P. J. Kretschmer, and C. F. Garon.** 1981. Molecular cloning of the K1 capsular polysaccharide genes of *E. coli*. Nature **289:**696–698.
142. **Silver, R. P., K. Prior, C. Nsahlai, and L. F. Wright.** 2001. ABC transporters and the export of capsular polysaccharides from gram-negative bacteria. Res. Microbiol. **152:**357–364.
143. **Smith, A. N., G. J. Boulnois, and I. S. Roberts.** 1990. Molecular analysis of the *Escherichia coli* K5 *kps* locus: identification and characterization of an inner-membrane capsular polysaccharide transport system. Mol. Microbiol. **4:**1863–1869.
144. **Smith, P. C., N. Karpowich, L. Millen, J. E. Moody, J. Rosen, P. J. Thomas, and J. F. Hunt.** 2002. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. Mol. Cell **10:**139–149.
145. **Steeghs, L., R. den Hartog, A. den Boer, B. Zomer, P. Roholl, and P. van der Ley.** 1998. Meningitis bacterium is viable without endotoxin. Nature **392:**449–450.
146. **Steenbergen, S. M., and E. R. Vimr.** 2008. Biosynthesis of the *Escherichia coli* K1 group 2 polysialic acid capsule occurs within a protected cytoplasmic compartment. Mol. Microbiol. **68:**1252–1267.
147. **Steenbergen, S. M., and E. R. Vimr.** 1990. Mechanism of polysialic acid chain elongation in *Escherichia coli* K1. Mol. Microbiol. **4:**603–611.
148. **Steenbergen, S. M., T. J. Wrona, and E. R. Vimr.** 1992. Functional analysis of the sialyltransferase complexes in *Escherichia coli* K1 and K92. J. Bacteriol. **174:**1099–1108.
149. **Steiner, K., R. Novotny, K. Patel, E. Vinogradov, C. Whitfield, M. A. Valvano, P. Messner, and C. Schaffer.** 2007. Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. J. Bacteriol. **189:**2590–2598.
150. **Steiner, K., R. Novotny, D. B. Werz, K. Zarschler, P. H. Seeberger, A. Hofinger, P. Kosma, C. Schaffer, and P. Messner.** 2008. Molecular basis of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus*. J. Biol. Chem. **283:**21120–21133.
151. **Stroeher, U. H., L. E. Karageorgos, R. Morona, and P. A. Manning.** 1992. Serotype conversion in *Vibrio cholerae* O1. Proc. Natl. Acad. Sci. U. S. A. **89:**2566–2570.
152. **Sugiyama, T., N. Kido, Y. Kato, N. Koide, T. Yoshida, and T. Yokochi.** 1997. Evolutionary relationship among *rfb* gene clusters synthesizing mannose homopolymer as O-specific polysaccharides in *Escherichia coli* and *Klebsiella*. Gene **198:**111–113.
153. **Sugiyama, T., N. Kido, Y. Kato, N. Koide, T. Yoshida, and T. Yokochi.** 1998. Generation of *Escherichia coli* O9a serotype, a subtype of *E. coli* O9, by transfer of the *wb** gene cluster of *Klebsiella* O3 into *E. coli* via recombination. J. Bacteriol. **180:**2775–2778.
154. **Sugiyama, T., N. Kido, T. Komatsu, M. Ohta, K. Jann, B. Jann, A. Saeki, and N. Kato.** 1994. Genetic analysis of *Escherichia coli* O9 *rfb*: identification and DNA sequence of phosphomannomutase and GDP-mannose pyrophosphorylase genes. Microbiology **140:**59–71.
155. **Suzuki, N., Y. Nakano, Y. Yoshida, H. Nakao, Y. Yamashita, and T. Koga.** 2000. Genetic analysis of the gene cluster for the synthesis of serotype a-specific polysaccharide antigen in *Actinobacillus actinomycetemcomitans*. Biochim. Biophys. Acta **1517:**135–138.
156. **Szabo, M., D. Bronner, and C. Whitfield.** 1995. Relationships between *rfb* gene clusters required for biosynthesis of identical D-galactose-containing O antigens in *Klebsiella pneumoniae* serotype O1 and *Serratia marcescens* serotype O16. J. Bacteriol. **177:**1544–1553.
157. **Szymanski, C. M., and B. W. Wren.** 2005. Protein glycosylation in bacterial mucosal pathogens. Nat. Rev. Microbiol. **3:**225–237.
158. **Tam, P. H., and T. L. Lowary.** 2009. Recent advances in mycobacterial cell wall glycan biosynthesis. Curr. Opin. Chem. Biol. **13:**618–625.
159. **Tang, G., and K. P. Mintz.** Glycosylation of the collagen adhesin EmaA of

*Aggregatibacter actinomycetemcomitans* is dependent upon the lipopolysaccharide biosynthetic pathway. J. Bacteriol. **192:**1395–1404.
160. **Tefsen, B., M. P. Bos, F. Beckers, J. Tommassen, and H. de Cock.** 2005. MsbA is not required for phospholipid transport in *Neisseria meningitidis*. J. Biol. Chem. **280:**35961–35966.
161. **Tzeng, Y. L., A. K. Datta, C. A. Strole, M. A. Lobritz, R. W. Carlson, and D. S. Stephens.** 2005. Translocation and surface expression of lipidated serogroup B capsular polysaccharide in *Neisseria meningitidis*. Infect. Immun. **73:**1491–1505.
162. **Vimr, E. R., and S. M. Steenbergen.** 2009. Early molecular-recognition events in the synthesis and export of group 2 capsular polysaccharides. Microbiology **155:**9–15.
163. **Vinogradov, E., E. Frirdich, L. L. MacLean, M. B. Perry, B. O. Petersen, J. O. Duus, and C. Whitfield.** 2002. Structures of lipopolysaccharides from *Klebsiella pneumoniae*. Elucidation of the structure of the linkage region between core and polysaccharide O chain and identification of the residues at the nonreducing termini of the O chains. J. Biol. Chem. **277:**25070–25081.
164. **Vionnet, J., and W. F. Vann.** 2007. Successive glycosyltransfer of sialic acid by *Escherichia coli* K92 polysialyltransferase in elongation of oligosialic acceptors. Glycobiology **17:**735–743.
165. **Wacker, M., M. F. Feldman, N. Callewaert, M. Kowarik, B. R. Clarke, N. L. Pohl, M. Hernandez, E. D. Vines, M. A. Valvano, C. Whitfield, and M. Aebi.** 2006. Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. Proc. Natl. Acad. Sci. U. S. A. **103:**7088–7093.
166. **Wacker, M., D. Linton, P. G. Hitchen, M. Nita-Lazar, S. M. Haslam, S. J. North, M. Panico, H. R. Morris, A. Dell, B. W. Wren, and M. Aebi.** 2002. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. Science **298:**1790–1793.
167. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945–951.
168. **Ward, A., C. L. Reyes, J. Yu, C. B. Roth, and G. Chang.** 2007. Flexibility in the ABC transporter MsbA: alternating access with a twist. Proc. Natl. Acad. Sci. U. S. A. **104:**19005–19010.
169. **Ward, C. K., and T. J. Inzana.** 1997. Identification and characterization of a DNA region involved in the export of capsular polysaccharide by *Actinobacillus pleuropneumoniae* serotype 5a. Infect. Immun. **65:**2491–2496.
170. **Ward, J. B.** 1981. Teichoic and teichuronic acids: biosynthesis, assembly, and location. Microbiol. Rev. **45:**211–243.
171. **Weigel, P. H., and P. L. DeAngelis.** 2007. Hyaluronan synthases: a decade-plus of novel glycosyltransferases. J. Biol. Chem. **282:**36777–36781.
172. **Whisstock, J. C., and A. M. Lesk.** 1999. SH3 domains in prokaryotes. Trends Biochem. Sci. **24:**132–133.
173. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu. Rev. Biochem. **75:**39–68.
174. **Whitfield, C., M. B. Perry, L. L. MacLean, and S. H. Yu.** 1992. Structural analysis of the O-antigen side chain polysaccharides in the lipopolysaccharides of *Klebsiella* serotypes O2(2a), O2(2a,2b), and O2(2a,2c). J. Bacteriol. **174:**4913–4919.
175. **Whitfield, C., J. C. Richards, M. B. Perry, B. R. Clarke, and L. L. MacLean.** 1991. Expression of two structurally distinct D-galactan O antigens in the lipopolysaccharide of *Klebsiella pneumoniae* serotype O1. J. Bacteriol. **173:**1420–1431.
176. **Yan, A., Z. Guan, and C. R. Raetz.** 2007. An undecaprenyl phosphate-aminoarabinose flippase required for polymyxin resistance in *Escherichia coli*. J. Biol. Chem. **282:**36077–36089.
177. **Yoshida, Y., Y. Nakano, Y. Yamashita, and T. Koga.** 1998. Identification of a genetic locus essential for serotype b-specific antigen synthesis in *Actinobacillus actinomycetemcomitans*. Infect. Immun. **66:**107–114.
178. **Young, N. M., J. R. Brisson, J. Kelly, D. C. Watson, L. Tessier, P. H. Lanthier, H. C. Jarrell, N. Cadotte, F. St. Michael, E. Aberg, and C. M. Szymanski.** 2002. Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. J. Biol. Chem. **277:**42530–42539.
179. **Zaitseva, J., S. Jenewein, T. Jumpertz, I. B. Holland, and L. Schmitt.** 2005. H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. EMBO J. **24:**1901–1910.
180. **Zayni, S., K. Steiner, A. Pfostl, A. Hofinger, P. Kosma, C. Schaffer, and P. Messner.** 2007. The dTDP-4-dehydro-6-deoxyglucose reductase encoding *fcd* gene is part of the surface layer glycoprotein glycosylation gene cluster of *Geobacillus tepidamans* GS5-97T. Glycobiology **17:**433–443.
181. **Zdorovenko, E. L., V. V. Ovod, G. V. Zatonsky, A. S. Shashkov, N. A. Kocharova, and Y. A. Knirel.** 2001. Location of the O-methyl groups in the O polysaccharide of *Pseudomonas syringae* pv. *phaseolicola*. Carbohydr. Res. **330:**505–510.
182. **Zhang, L., A. al-Hendy, P. Toivanen, and M. Skurnik.** 1993. Genetic organization and sequence of the rfb gene cluster of *Yersinia enterocolitica* serotype O:3: similarities to the dTDP-L-rhamnose biosynthesis pathway of *Salmonella* and to the bacterial polysaccharide transport systems. Mol. Microbiol. **9:**309–321.